

# Ethical issues associated with HIV phylogenetics in HIV transmission dynamics research: A review of the literature using the Emanuel Framework

Farirai Mutenherwa  | Douglas R. Wassenaar | Tulio de Oliveira

## Correspondence

Farirai Mutenherwa, University of KwaZulu-Natal, School of Applied Human Sciences, Pietermaritzburg, South Africa, and the KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa.  
Email: mutenherwaf@gmail.com

## Abstract

The reduced costs of DNA sequencing and the use of such data for HIV-1 clinical management and phylogenetic analysis have led to a massive increase of HIV-1 sequences in the last few years. Phylogenetic analysis has shed light on the origin, spread and characteristics of HIV-1 epidemics and outbreaks. Phylogenetic analysis is now also being used to advance our knowledge of the drivers of HIV-1 transmission in order to design effective interventions. However, HIV phylogenetic analysis presents unique ethical challenges, which have not been fully explored. This review presents an analysis of what appear to be key ethical issues in HIV phylogenetics in the hope of stimulating further conceptual and empirical work in this rapidly emerging area. We structure the review using the Emanuel Framework, a systematic, holistic framework, which has been adapted for use in developing countries, which bear the brunt of the HIV-1 pandemic.

## KEYWORDS

HIV phylogenetics, HIV transmission dynamics, Molecular epidemiology, Public health ethics, research ethics

## 1 | INTRODUCTION

Over the last decade, phylogenetic analysis has led to remarkable advances in understanding the origin, spread and characteristics of many infectious diseases and epidemics. For example, phylogenetic analysis of viral sequences was instrumental in confirming that Zika virus strains from the Pacific Islands were the source of the Zika epidemic in Brazil<sup>1</sup> and that the Ebola epidemic in West Africa was linked to an initial outbreak in Sierra Leone.<sup>2</sup> These techniques were also used to identify the origin of the worldwide HIV epidemic from strains that existed in West-Central Africa in the early 1900s. Phylogenetic analysis is currently being applied to identify the

underlying drivers of HIV-1 transmission at a population level across the globe.<sup>3</sup>

Phylogenetic analysis provides an optimistic outlook for deriving interventions to reduce HIV-1 transmission. However, it also raises complex ethical issues that could hinder its success. In this review, we outline specific applications of HIV-1 phylogenetic analysis in relation to an ethical framework that can be applied to HIV phylogenetics research, the Emanuel Framework (EF). The EF framework comprises eight broad principles: 1) community participation, 2) social value, 3) scientific validity, 4) fair participant selection, 5) favourable risk-benefit ratio, 6) independent ethics review, 7) informed consent, and 8) on-going respect for participants.

<sup>1</sup>Faria NR, et al. Zika Virus in the Americas: Early Epidemiological and Genetic Findings. *Science* 2016; 352: 345-9.

<sup>2</sup>Park DJ, et al. Ebola Virus Epidemiology, Transmission, and Evolution During Seven Months in Sierra Leone. *Cell* 2015; 161: 1516-26.

<sup>3</sup>Ratmann O, et al. Sources of HIV Infection Among Men Having Sex with Men and Implications for Prevention. *Sci Transl Med* 2016; 8: 320; de Oliveira T, et al. Transmission Networks and Risk of HIV Infection in KwaZulu-Natal, South Africa: a Community-wide Phylogenetic Study. *Lancet HIV* 2017; 4: e41-e50.

## 2 | HIV-1 PHYLOGENETIC ANALYSIS

Phylogenetic analysis is a scientific process used to inspect small disparities in viral genes using computational techniques in order to determine the genetic distance between different strains.<sup>4</sup> The process begins with the generation of viral sequences, which are primarily derived from virions isolated from blood samples of HIV-infected persons. A detailed review of phylogenetic methods is documented elsewhere<sup>5</sup> and is beyond the scope of this review. However, some applications of HIV phylogenetic analysis are summarised below.

Phylogenetic analysis is commonly used to identify the potential source of HIV-1 transmission events.<sup>6</sup> This allows scientists to rule out or confirm a possible specific partner or contact as the source of HIV infection. The analysis normally includes sequences of the suspected transmission case in addition to control datasets, which can be extracted from a genetic database and/or generated from other individuals infected with HIV in the same community.<sup>7</sup> If the two sequences are more closely related by genetic distance (the extent to which they have diverged from a common ancestor) to each other than they are to the comparison samples, and if the relationship was not by chance, it is assumed that the two individuals are in a linked transmission chain.

A classic example of the use of phylogenetic analysis to resolve a transmission case can be seen in the work of Goedhals and colleagues<sup>8</sup> who identified the source of a breastfed surrogate HIV-1 transmission event. In this case, researchers generated HIV-1 sequence data from the suspected transmission case (baby-case and his aunt and cousin) and analysed it together with control sequences, which had been collected in the local community. They then applied three methods for the construction of phylogenies (NJ, ML and Bayesian). These methods all showed that the sequences from the suspected transmission case were closely related based on statistical tests and estimates. The phylogenetic analysis supported interviews between public health officials and family members, which revealed that the aunt, who was HIV-1 positive, breastfed the baby-case when the baby's mother went to work. Even though a large control dataset of sequences was used, the authors highlighted that phylogenetic analysis alone could not solve the case.

<sup>4</sup>Bernard EJ, et al. HIV Forensics: Pitfalls and Acceptable Standards in the Use of Phylogenetic Analysis as Evidence in Criminal Investigations of HIV Transmission. *HIV Med* 2007; 8: 382-387.

<sup>5</sup>Ibid; Yang Z, Rannala B. Molecular Phylogenetics: Principles and Practice. *Nat Rev Genet* 2012; 13: 303-314.

<sup>6</sup>Liu SH et al. Viral Genetic Linkage Analysis in the Presence of Missing Data. *PLoS One* 2015; 10: e0135469; Grabowski, *op. cit.* note 5, p. 127-129.

<sup>7</sup>Scaduto DI et al. Source Identification in Two Criminal Cases Using Phylogenetic Analysis of HIV-1 DNA Sequences. *Proc Natl Acad Sci U S A* 2010; 107: 21242-21247; UNAIDS. 2013. *Ending overly broad criminalization of HIV non-disclosure, exposure and transmission: Critical scientific, medical and legal considerations*. Geneva. Available at: [http://www.unaids.org/sites/default/files/media\\_asset/20130530\\_Guidance\\_Ending\\_Criminalisation\\_0.pdf](http://www.unaids.org/sites/default/files/media_asset/20130530_Guidance_Ending_Criminalisation_0.pdf) [Accessed 4 July 2017].

<sup>8</sup>Goedhals D et al. The Tainted Milk of Human Kindness. *Lancet* 2012; 380: 702.

HIV phylogenetics can also be used in prevention trials. A good example was the HPTN-052 trial, which was designed to evaluate antiretroviral and other HIV interventions among HIV sero-discordant couples. In the trial, HIV-1 phylogenetic analysis was used to establish whether the HIV positive partner infected their HIV-negative partner. Results in this trial showed that 18.4% of new HIV infections in sero-discordant couples did not originate from their current primary partner.<sup>9</sup> Phylogenetic analysis was used with viral load clinical measurements and interviews to calculate the effectiveness of the trial. Another example of the use of HIV-1 phylogenetics was in the Partners in Prevention (PIP) trial, which estimated that 26.5% of new HIV infections were unlinked to the index HIV-positive partner.<sup>10</sup>

Phylogenetic analysis can also be used for epidemiological purposes and is increasingly applied in studies that seek to understand population-based HIV transmission patterns and dynamics. For example, a large-scale phylogenetic study of HIV genetic sequences from men who have sex with men (MSM) in the Netherlands was used to determine the drivers of the underlying high-level transmission of HIV-1 in this population.<sup>11</sup> It was estimated that 71% of transmissions were from undiagnosed men and that 43% of the transmissions occurred in the first year of HIV-1 infection. The study also assessed the effectiveness of interventions to reduce HIV-1 transmission and found that increased annual testing coverage, pre-exposure ARV prophylaxis and immediate treatment were likely to avert 75% of new infections.

Routine HIV-1 sequence data has enabled public health practitioners to monitor HIV transmission hotspots in order to guide public health responses in near real time. Poon and colleagues describe a Canadian implementation case study where an automated phylogenetic system monitored a clinical database to prevent an HIV outbreak.<sup>12</sup> According to local treatment guidelines, all HIV positive patients were required to undergo routine HIV genotype testing at time of diagnosis. Within three months, a growing cluster of young MSMs was identified. In that cluster, eleven new cases were detected, eight of whom were infected by drug resistant HIV strains. The results prompted a public health follow-up on the affected subpopulation to ensure that members were linked to treatment and care. The public health intervention reduced drug resistant HIV transmission within the subpopulation. Similar work was also conducted in the USA, where HIV-1 transmission networks were reconstructed with a view to

<sup>9</sup>Eshleman SH, Hudelson SE, Redd AD. Analysis of Genetic Linkage of HIV from Couples Enrolled in The HIV Prevention Trials Network 052 Trial. *J. Infect. Dis.* 2011; 204: 1918 - 1926.

<sup>10</sup>Campbell MS et al. Viral Linkage in HIV-1 Seroconverters and their Partners in an HIV-1 Prevention Clinical Trial. *PLoS One* 2011; 6: e16986.

<sup>11</sup>Ratmann, *op. cit.* note 3, pp. 3-5.

<sup>12</sup>Poon AF et al. Near Real-time Monitoring of HIV Transmission Hotspots from Routine HIV Genotyping: An Implementation Case Study. *Lancet HIV* 2016; 3: e231-e238.

assessing and guiding the prevention of HIV transmission in the community.<sup>13</sup>

Phylogenetic techniques are increasingly used to understand HIV-1 transmission in Africa. Findings from a phylogenetic study in Rakai, Uganda, showed that 44% of HIV infections were transmitted in stable household partnerships. Of those transmissions that occurred outside the household partnership, 62% were from individuals from another community.<sup>14</sup> Results from a recent community-wide phylogenetic study in South Africa showed that the majority of transmissions to young women (<25 years) were likely to be from men approximately 8.7 years older than them.<sup>15</sup>

Lastly, HIV-1 phylogenetic analysis has also been widely used in forensic work. This has predominantly been applied in countries where HIV transmission is criminalized.<sup>16</sup> As Abecasis and colleagues<sup>17</sup> argue, the use of phylogenetics as evidence in court should be seen in the context of hypothesis testing. A normal court null hypothesis is that the defendant infected the victim. In such cases, phylogenetic analysis and expert testimony are used to provide evidence that supports or refutes the null hypothesis. While phylogenetics can sometimes prove innocence, it is not possible to use phylogenetics alone to make a guilty verdict<sup>18</sup>. Furthermore, phylogenetic analysis alone cannot provide definitive proof of the route, direction and timing of HIV transmission between two people. There may be other possible reasons why the two individuals have similar viruses. For example, a third person may exist, who is the original source of the transmission event, but who was not sampled.<sup>19</sup>

Phylogenetic analysis is most informative when HIV sequences are linked to clinical, demographic and behavioural data of the sampled individuals. Sources of this data could be medical records and surveys, which contain demographic, clinical and sexual behaviour information. The source of the genetic material used in HIV-1 phylogenetic analysis is the virus itself, not human genes. For this reason, the ethical issues traditionally linked to human genetic research, such as privacy and confidentiality, fears of stigma and discrimination, may, at first glance, appear to be of little concern. A closer examination, however, reveals distinct ethical considerations.

<sup>13</sup>Little SJ et al. Using HIV Networks to Inform Real Time Prevention Interventions. *PLoS One* 2014; 9: e98443.

<sup>14</sup>Grabowski MK et al. The Role of Viral Introductions in Sustaining Community-based HIV Epidemics in Rural Uganda: Evidence from Spatial Clustering, Phylogenetics, and Egocentric Transmission Models. *PLoS Med* 2014; 11: e1001610.

<sup>15</sup>de Oliveira T et al., *op. cit.* note 3, pp. e41-e50.

<sup>16</sup>Scaduto et al., *op. cit.* note 8, pp. 21242-21247; Ou CY et al. Molecular Epidemiology of HIV Transmission in a Dental Practice. *Science* 1992; 256: 1165-1171.

<sup>17</sup>Abecasis AB et al. Science in Court: The Myth of HIV Fingerprinting. *Lancet Infect Dis* 2011; 11: 78-79.

<sup>18</sup>Ibid.

<sup>19</sup>Ibid; Bernard, *op. cit.* note 4, p. 385.

### 3 | AN ETHICAL FRAMEWORK FOR THE REVIEW OF HIV PHYLOGENETICS RESEARCH: THE EMANUEL, WENDLER AND GRADY FRAMEWORK

In order to systematically review and organise the literature, we used a frequently cited<sup>20</sup> ethics framework developed by Emanuel, Wendler and Grady.<sup>21</sup> The framework is referred to as the Emanuel Framework (EF) throughout this review. We chose the EF primarily on merit. The EF was developed from content analysis of several major international normative ethics research guidelines. Although initially developed for clinical research in developed countries, the EF has since been adapted for use in developing countries<sup>22</sup> and for the review of other types of research, notably, social science<sup>23</sup> and health systems research.<sup>24</sup> Additional uses of the EF are documented elsewhere.<sup>25</sup>

Because we assumed that the EF could usefully accommodate most of the ethical issues in HIV phylogenetic research, the EF is outlined in detail below, followed by an identification of the key ethical issues in HIV phylogenetic research using the same framework.

#### 3.1 | Community participation

Ethical research requires the development of collaborative partnerships between researchers and the target community. The engagement should take place throughout the research process from conceptualisation and implementation to dissemination of results in order to ensure that the interests of the community are considered at each stage. Such measures are meant to prevent or minimise exploitation of research participants.

A conceptual issue that may create problems with community engagement efforts for phylogenetic research is that communities are often not stable, static entities with clearly defined boundaries.

<sup>20</sup>Wassenaar D, Rattani A. What Makes Health Systems Research in Developing Countries Ethical? Application of the Emanuel Framework for Clinical Research to Health Systems Research. *Developing World Bioethics* 2016; 16: 133-139; Tsoka-Gwegweni J, Wassenaar DR. Using the Emanuel framework to examine ethical issues raised by a biomedical research ethics committee in South Africa. *Journal of Empirical Research on Human Research Ethics* 2014; 9: 36-45.

<sup>21</sup>Emanuel EJ, Wendler D, Grady C. What Makes Clinical Research Ethical? *JAMA* 2000; 283: 2701-2711.

<sup>22</sup>Emanuel EJ, Wendler D, Killen J, Grady C. What Makes Clinical Research in Developing Countries Ethical? The Benchmarks of Ethical Research. *J Infect Dis* 2004; 189: 930-937; Tsoka-Gwegweni J, Wassenaar DR. Using the Emanuel Framework to Examine Ethical Issues Raised by a Biomedical Research Ethics Committee in South Africa. *J Empir Res Hum Res Ethics* 2014; 9: 36-45.

<sup>23</sup>Wassenaar DR, Mamotte N. 2012. Ethical Issues and Ethics Reviews in Social Science Research. In: *Oxford Handbook of International Psychological Ethics*. Leach MM, Stevens MJ, Lindsay G, Ferrero A, Korkut Y, editors: Oxford University Press; Mutenherwa F, Wassenaar D. 2014. Ethics Review of Social and Behavioural Research in an African Context. In: *Research Ethics in Africa: A Resource for Research Ethics Committees*. Kruger M, Ndebele P, Horn L, editors: SUN Media Stellenbosch.

<sup>24</sup>Wassenaar D, Rattani A. What Makes Health Systems Research in Developing Countries Ethical? Application of the Emanuel Framework for Clinical Research to Health Systems Research. *Dev World Bioeth* 2016; 16: 133-139.

<sup>25</sup>Tsoka-Gwegweni, Wassenaar, *op. cit.* note 21, pp. 37-38.

Rather, they are fluid and may consist of sub-communities whose members may not share similar research needs and priorities. It may, therefore, be difficult to identify legitimate community members and stakeholders to help plan and conduct the study and disseminate the results. Even more challenging could be the identification of representatives of relevant communities who have the best interests of the target community at heart.

Evaluating the success and adequacy of community engagement may also be difficult, particularly in the absence of clearly defined matrices for its assessment.<sup>26</sup> Similarly, considering the number of players in the engagement process in a given community, there should be consensus on who legitimises the community engagement process. Such issues are not always straightforward as different stakeholders and levels of authority in a community might possess conflicting views.

HIV phylogenetic research can have a positive (for example, better prevention and treatment programmes) or negative (for example, stigma associated with HIV transmission for sub-groups identified as high transmitters) impact on communities. For this reason and due to its complexity, investigators need to invest heavily in community engagement efforts<sup>27</sup> in order to ensure that research messages are appropriately packaged for target audiences. Inappropriate messaging could undermine the validity of informed consent, exposing communities and research participants to risk of exploitation.

### 3.2 | Social value

Research should address socially valuable questions in order to justify involving human participants.<sup>28</sup> An example of a valuable question is one that generates new knowledge or understanding on human health or illness. It is therefore critical that in developing research projects, researchers should identify the study beneficiaries and outline the benefits that will accrue to them. Benefits can be classified as direct or indirect and can occur immediately or at a later date.<sup>29</sup> It is also important to develop mechanisms that enhance the value of research and to consider the potential impact of the research on the existing health-care infrastructure and social system.<sup>30</sup>

HIV phylogenetic research can potentially benefit both HIV positive and negative individuals, communities affected with HIV and AIDS, health systems and society at large. Comparatively, its contribution could be more pronounced in sub-Saharan Africa where the HIV burden is greatest. When phylogenetic analysis of HIV

sequences is used in conjunction with detailed epidemiological, clinical, demographic and behavioural data, it generates rich information on HIV transmission dynamics at community, regional and country levels.<sup>31</sup> Specifically, phylogenetic studies of HIV transmission can identify key traits associated with individuals or sub-populations responsible for onward transmission of HIV and use this information to derive effective interventions. For example, phylogenetic research can be used to identify HIV transmission geographical hotspots, the characteristics of individuals responsible for a disproportionately large number of infections and the local source of HIV drug resistant strains. A clear understanding of the patterns and determinants of HIV transmission and drug resistance acquisition is critical for the identification and design of the most effective HIV prevention and treatment approaches.<sup>32</sup>

Disseminating research results to key stakeholders can enhance the social value of HIV phylogenetic studies. Key stakeholders for HIV phylogenetic research can be identified through collaborative partnerships and may include people living with HIV, advocacy groups, clinicians, health policy makers, developmental agencies and public health officials. These stakeholders play an instrumental role in translating research into policy and should be engaged throughout the research process to enhance the social value of the study. Apart from journal publications, conference presentations and policy briefs, community feedback meetings could also be organised to ensure that results reach grassroots levels and that the views of the community are taken into consideration. Another approach that may enhance social value is to integrate phylogenetic research into long-term health strategies and/or public health programmes.

Another important aspect of social value is to assess the impact of the research on health-seeking behaviour. For example, HIV phylogenetic research can be seen as a threat to individual privacy as HIV phylogenies can potentially identify individuals that are linked to population groups that are stigmatised or penalized, such as MSM or sex workers. These individuals may avoid positive health-seeking behaviours that could benefit themselves and the population out of fear of stigma or prosecution. For example, they may refrain from HIV diagnosis and drug resistance testing and this may lead to new infections and continued transmission of HIV drug resistant strains.<sup>33</sup> Similarly, individuals may avoid participating in studies that use HIV genetic data. The social value of HIV phylogenetic research is further discussed in this review under the ethical principle of Favourable risk-benefit ratio.

### 3.3 | Scientific validity

The study design should be rigorous enough to ensure that valid, reliable, interpretable and in some cases generalizable data is

<sup>26</sup>MacQueen KM et al. Evaluating Community Engagement in Global Health Research: The Need for Metrics. *BMC Med Ethics* 2015; 16: 44.

<sup>27</sup>UNAIDS-AVAC. 2011. *Good participatory practice: Guidelines for biomedical HIV prevention trials*. Geneva: UNAIDS-AVAC. Available at: [http://www.unaids.org/sites/default/files/media\\_asset/JC1853\\_GPP\\_Guidelines\\_2011\\_en\\_0.pdf](http://www.unaids.org/sites/default/files/media_asset/JC1853_GPP_Guidelines_2011_en_0.pdf) [Accessed 10 July 2017].

<sup>28</sup>Emanuel EJ, Wendler D, Grady C. 2008. An Ethical Framework for Biomedical Research. In: *The Oxford textbook of clinical research ethics* Emanuel EJ, Grady C, Crouch RA, Lie RK, Miller FG, Wendler D, editors. New York: Oxford University Press.

<sup>29</sup>Wendler D, Rid A. In Defense of a Social Value Requirement for Clinical Research. *Bioethics* 2017; 31: 77-86.

<sup>30</sup>Emanuel et al., *op. cit.* note 27, p. 127.

<sup>31</sup>Dennis AM et al. Phylogenetic Studies of Transmission Dynamics in Generalized HIV Epidemics: An Essential Tool Where the Burden is Greatest? *JAIDS* 2014; 67: 181-195.

<sup>32</sup>Smith DM et al. A Public Health Model for the Molecular Surveillance of HIV Transmission in San Diego, California. *AIDS* 2009; 23: 225-232; Poon AF et al., *op. cit.* note 12, p. e231-e238; Little et al., *op. cit.* note 13, p. e98443.

<sup>33</sup>Brooks JI, Sandstrom PA. The Power and Pitfalls of HIV Phylogenetics in Public Health. *Can J Public Health* 2013; 104: e348-e350.

generated. Scientific rigor and integrity are key elements of ethical research. Poor quality science is normally produced by the use of unreliable and/or invalid research methods. It also constitutes unethical conduct because if results from such studies are not trustworthy, resources are wasted and research participants are exposed to potential risk of harm and inconvenience for no apparent benefit.<sup>34</sup> 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/IRB) reviews and approves the protocol.

HIV phylogenetics currently has methodological shortcomings, which require expert consideration to avoid or minimise erroneous interpretations and conclusions. These include: (A) the inability of the technique to establish the direction of transmission, (B) the effects of co-infection and super-infection, (C) the lack of standard phylogenetic cut-offs for the identification of transmission clusters, (D) the variability of sampling coverage (there are normally missing individuals) and lastly, (E) the absence of appropriate control sequences. Item (A) to (C) are discussed in the next paragraphs (3.3.1 to 3.3.3). However, the last two items, (D) and (E), are addressed under the fourth EF principle: 'Fair participant selection' (as 3.4.1 and 3.4.2).

### 3.3.1 | Phylogenetics and direction of transmission

One of the key limitations of HIV phylogenetic analysis is its inability to provide definitive proof of direction of HIV transmission between two people on its own. This lack of certainty and precision has raised concerns about the reliability of phylogenetic analysis in reconstructing the HIV transmission history between two or more individuals in a phylogenetic cluster.<sup>35</sup> In HIV forensics, these problems potentially result in miscarriages of justice through erroneous convictions or acquittals, hence the need to interpret results with extreme caution.<sup>36</sup> In the absence of specific evidence on the source of transmission, those considered undesirable in society may face prosecution.<sup>37</sup> Apart from criminal prosecution, phylogenetics analysis could also have serious implications for the course of justice in civil cases.

### 3.3.2 | Effects of co-infection and super infection

The difficulty in identifying the precise source of HIV is compounded by the fact that HIV infection can be caused by more than one strain that is constantly mutating, recombining and evolving into different

strains, even within the same individual.<sup>38</sup> In addition, over time, individuals may be infected with different HIV strains that have genetically distinctive characteristics.<sup>39</sup> Certain viral sequences, especially those generated from direct Sanger sequencing techniques, may not fully represent the viral diversity within an individual.<sup>40</sup> This may bias the analysis and have serious consequences.

### 3.3.3 | Lack of standard phylogenetic cut-offs for the identification of transmission clusters

An HIV cluster is a group of sequences that are more similar to each other than to other sequences in the same dataset, based on a pre-defined criterion or algorithm. However, there is no gold standard measure for clustering, hence the definition of a cluster remains largely subjective.<sup>41</sup> Furthermore, although several genetic clustering methods and related computer software are widely available, most of these methods have neither been validated on already known clusters nor evaluated using the same dataset. The interpretation of HIV genetic clusters may, therefore, be ambiguous and biased, which could lead to misplaced priorities for HIV interventions. As one author observed, due to the methodological issues highlighted, "... the research community needs to have greater skepticism about clustering methods and, ultimately, to reach a consensus on best practices for generating and interpreting clusters."<sup>42</sup> Nonetheless, the application of clustering methods is common, with opportunities for improvement. A review of the impact and shortcomings of clustering methods is documented elsewhere<sup>43</sup> and is beyond the scope of this paper.

### 3.4 | Fair participant selection

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 unethical. Those who are selected to participate in the study should also be informed of the research results and receive any benefit that comes from them. In the next paragraphs, we will discuss the two remaining

<sup>34</sup>Emanuel et al., *op. cit.* note 21, p. 933.

<sup>35</sup>Pillay D et al. HIV Phylogenetics. *BMJ* 2007; 335: 460-461; Bernard et al., *op. cit.* note 4, pp. 384-385.

<sup>36</sup>Metzker ML et al. Molecular Evidence of HIV-1 Transmission in a Criminal Case. *Proc Natl Acad Sci U S A* 2002; 99: 14292-14297.

<sup>37</sup>Willyard C. Africa's HIV Transmission Laws Based on Questionable Science. *Nat Med* 2007; 13: 890.

<sup>38</sup>Abecasis et al., *op. cit.* note 17, pp. 78-79.

<sup>39</sup>Blick G et al. The Probable Source of Both the Primary Multidrug-resistant (MDR) HIV-1 Strain Found in a Patient with Rapid Progression to AIDS and a Second Recombinant MDR Strain Found in a Chronically HIV-1-infected Patient. *J Infect Dis* 2007; 195: 1250-1259; Redd AD, Quinn TC, Tobian AA. Frequency and Implications of HIV Superinfection. *Lancet Infect Dis* 2013; 13: 622-628; Hecht FM, Wolf LE, Lo B. Lessons from an HIV Transmission Pair. *J Infect Dis* 2007; 195: 1239-1241.

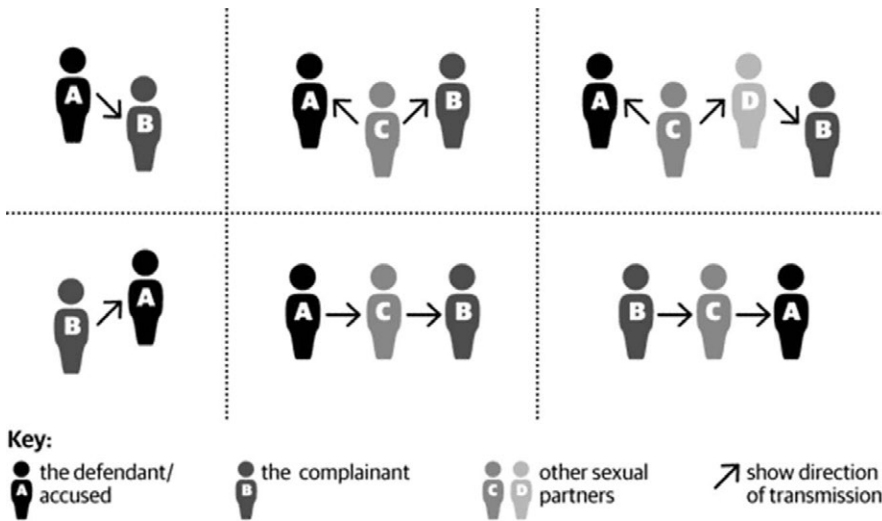
<sup>40</sup>Grabowski, Redd, *op. cit.* note 5, pp. 127-128.

<sup>41</sup>Poon AF. Impacts and shortcomings of genetic clustering methods for infectious disease outbreaks. *Virus evolution* 2016; 2: vew031; Hassan AS, Pybus OG, Sanders EJ, Albert J, Esbjornsson J. Defining HIV-1 transmission clusters based on sequence data: a systematic review and perspectives. *AIDS* 2017.

<sup>42</sup>Poon, *op. cit.* note 41, p. 8.

<sup>43</sup>Ibid.





**FIGURE 1** Potential routes and directions of HIV transmission.

methodological shortcomings: (D) effect of HIV-1 sampling coverage and (E) use of appropriate controls.

### 3.4.1 | Sampling coverage

It is virtually impossible to include all potential sources of HIV transmission for phylogenetic analysis because some members in the transmission network may not be available for sampling. Reasons for non-availability could include migration, mortality, informed and uninformed refusals, viral suppression and other practical challenges associated with HIV-1 sample collection and sequencing. Phylogenetic analysis cannot, therefore, exclude an indirect link to HIV transmission unless all potential sources of HIV transmission are included in the sample for analysis. We illustrate the potential sources of HIV-1 transmission and the effect of sampling in Figure 1.<sup>44</sup> In this schematic case, there are two or more people, A, B, C and D. A is suspected of infecting B with HIV. Given incomplete sampling, several possible explanations could be given to account for B's infection: (i) A could have directly passed on the virus to B; (ii) B could have directly passed on the virus to A; (iii) There could be an unknown third party, C, who may not have been included in the sample but who could have infected both A and B; (iv) A could have infected C, who in turn passed on the virus to B or (v) C could have infected A and another unsampled individual, D, who could subsequently have infected B.

### 3.4.2 | Use of appropriate control samples

To enhance the validity and reliability of phylogenetic analysis, appropriate control samples must be used. These controls should come from the same viral subtype and geographic region and should ideally be collected and sequenced at the same time. Strict laboratory

protocols and standards should also be followed.<sup>45</sup> Such precautions are especially important for forensic cases or complex cases of HIV transmission. For example, in the infamous Lafayette case, a gastroenterologist was convicted of attempted second-degree murder for injecting his former girlfriend with HIV-1. In their evidence, the researchers needed to generate control datasets for HIV-1 positive individuals from the same area and use two distinct laboratories with strict control protocols.<sup>46</sup> Another example already mentioned, was the use of a large control dataset of thousands of sequences to solve the HIV-1 transmission surrogate case in South Africa.<sup>47</sup> In the study of the MSM epidemic in the Netherlands, large datasets of HIV-1 control sequences from Europe were needed to identify transmission clusters.<sup>48</sup>

### 3.5 | Favourable risk-benefit ratio

A favourable risk-benefit ratio is realised when research benefits and burdens are fairly distributed.<sup>49</sup> Research benefits could apply to research participants, to society in general and/or to the study's contribution to scientific knowledge. Burdens could be social, physical, psychological, economic or legal. Such factors should all be considered in making risk benefit assessments. If potential harms are identified, researchers need to take appropriate measures to prevent or minimise them. In addition, researchers should try to enhance benefits to create a favourable risk-benefit ratio.<sup>50</sup> Under no circumstances should payment for participation in research be used to offset research risks and burdens. In assessing

<sup>44</sup>NAM AIDS MAP. 2017. *Proving a Cause-effect Relationship Between the Defendant's Behaviour and the Alleged Outcome*. London: NAM Publications. Available at: <http://www.aidsmap.com/Proving-a-cause-effect-relationship-between-the-defendants-behaviour-and-the-alleged-outcome/page/1444124/> [Accessed 11 July 2017].

<sup>45</sup>Bernard EJ et al. 2007. *HIV Forensics: The Use of Phylogenetic Analysis as Evidence in Criminal Investigation of HIV Transmission*. National AIDS Manual Briefing Papers, February 2007. Available at: <https://ssrn.com/abstract=967915> [Accessed 29 June 2017].

<sup>46</sup>Metzker et al., *op. cit.* note 35, p. 14293.

<sup>47</sup>Goedhals et al., *op. cit.* note 9, p.702.

<sup>48</sup>Ratmann et al., *op. cit.* note 3, pp. 3-8.

<sup>49</sup>Emanuel et al., *op. cit.* note 20, p. 2705.

<sup>50</sup>Ibid.

risks, it is critical to consider the probability of harm occurring, as well as its severity.

The benefits of data sharing to advance scientific progress and public health are widely acknowledged.<sup>51</sup> Sharing of sequence datasets can help formulate and address new research questions, inform the design and implementation of future epidemiological studies and provide a unique opportunity for meta-analyses of available data, which can be used to track and predict future epidemics.<sup>52</sup> Access to datasets enables other researchers to critically evaluate published results, which might foster greater integrity among researchers. In addition, this minimises the costs of generating additional data. However, there is a tension between the public health benefits of data sharing and the risk to privacy for individuals whose data is used in HIV genetic research.

In spite of the increased accessibility of genomic sequence data there are no adequate laws, regulations and guidelines that protect individual privacy.<sup>53</sup> For example, a global privacy governance framework for genomic databases, and by extension genomic research, is non-existent.<sup>54</sup> Furthermore, where available, national guidelines on privacy are often inadequate, fragmented, diverse and complicated. This was illustrated by results of an analysis of laws from twenty developed and developing countries.<sup>55</sup> Although the reviewed legislation and frameworks were conducted with biobanks in mind, similar concerns may also arise in phylogenetic research because of its reliance on genomic databases. The existing ethical and legal framework could therefore hinder collaborative research across national borders. Another challenge is the absence of a clear definition of what the true benefit or actual risk is in the context of HIV molecular epidemiology research.<sup>56</sup>

In their seminal host genetics study, Gymrek and colleagues demonstrated that individuals who participated in a genomic study could be identified using free publicly available information even though their genetic and personal information was stored in databases in de-identified form.<sup>57</sup> In the consent documents for the study, the risk of re-identification was only mentioned as a distant possibility. This implied that the researchers and possibly the research participants did not fully understand the risks to privacy at

the onset of the genetic study. Other studies have also raised similar privacy concerns.<sup>58</sup>

The privacy threats mentioned above occurred in host genetic studies. However, similar violations could occur in HIV genetic studies because the viral sequence is traceable to its human host. For example, a person can have access to his own HIV drug resistance genotype and search public databases, allowing the identification of closely related sequences. Additionally, advances in genomic techniques like NGS and their increased availability enable scientists to draw more accurate conclusions about HIV transmission between individuals.<sup>59</sup> Such advances could provide robust information that may lead to loss of privacy.<sup>60</sup> Privacy violations are compounded by the increasing and often complex interface between clinical management and use of less regulated Internet based platforms where personal and health information can be available.

Privacy concerns related to HIV genetic sequences become pronounced when HIV phylogenetic techniques use public data. In HIV phylogenetics, the sequences can be ordinarily generated from patients during HIV drug resistance testing during routine clinical patient management. The sequences are then used, together with public data, to examine and infer putative local or global HIV transmission clusters. Considering that HIV genetic data is arguably host-specific and each sequence is almost unique for each infected person<sup>61</sup> it is possible to identify individuals with sequences that have a high degree of similarity. These individuals would be seen as sharing a putative transmission link and possibly sharing the same source of infection. Putative linkages have serious legal (both civil and criminal) implications, especially in countries and states where HIV is criminalised<sup>62</sup> or where certain groups involved in HIV transmission are stigmatised or penalized.

Linking socio-demographic data with putative HIV transmissions can also lead to unintended recognition of subgroups or individuals with unique characteristics, for example, those thought to be associated with high-risk behaviours. Identification of such population subgroups can lead to increased social stigma. Examples of social groups that have historically suffered, and continue to suffer, social stigma includes injection drug users, commercial sex workers and

<sup>51</sup>S. Bull et al. Best Practices for Ethical Sharing of Individual-Level Health Research Data From Low- and Middle-Income Settings. *J Empir Res Hum Res Ethics* 2015; 10: 302-313.

<sup>52</sup>Kuiken C, Korber B, Shafer RW. HIV Sequence Databases. *AIDS Rev* 2003; 5: 52-61.

<sup>53</sup>Mehta SR, Vinterbo SA, Little SJ. Ensuring Privacy in the Study of Pathogen Genetics. *Lancet Infect Dis* 2014; 14: 773-777.

<sup>54</sup>Dove ES. Biobanks, Data Sharing, and the Drive for a Global Privacy Governance Framework. *J Law Med Ethics* 2015; 43: 675-89.

<sup>55</sup>The analyses covers the following countries and jurisdictions: Australia, Brazil, China, Denmark, France, India, Israel, Nigeria, South Africa, Spain, Taiwan, the United Kingdom, Canada, Estonia, the European Union, Finland, Germany, Japan, Mexico, the Netherlands, Uganda, and the United States. It was published in the *Journal of Law, Medicine & Ethics* Volume 43, Issue 4 and Volume 44, Issue 1. Under the title: Harmonizing Privacy Laws to Enable International Biobank Research.

<sup>56</sup>Mehta et al., *op. cit.* note 52, p. 773.

<sup>57</sup>Gymrek M et al. Identifying Personal Genomes by Surname Inference. *Science* 2013; 339: 321-324.

<sup>58</sup>Erich Y, Narayanan A. Routes for Breaching and Protecting Genetic Privacy. *Nat Rev Genet* 2014; 15: 409-421; Homer N et al. Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays. *PLoS Genet* 2008; 4: e1000167; Im HK et al. On Sharing Quantitative Trait GWAS Results in an Era of Multiple-omics Data and the Limits of Genomic Privacy. *Am J Hum Genet* 2012; 90: 591-598; L.L. Rodriguez et al. Research Ethics. The Complexities of Genomic Identifiability. *Science* 2013; 339: 275-276; Schadt EE, Woo S, Hao K. Bayesian Method to Predict Individual SNP Genotypes from Gene Expression Data. *Nat Genet* 2012; 44: 603-608.

<sup>59</sup>Brumme ZL et al. HLA-associated Immune Escape Pathways in HIV-1 Subtype B Gag, Pol and Nef Proteins. *PLoS One* 2009; 4: e6687; Poon AF et al. Dates of HIV Infection Can be Estimated for Seroprevalent Patients by Coalescent Analysis of Serial Next-generation Sequencing Data. *AIDS* 2011; 25: 2019-2026.

<sup>60</sup>Brooks, Sandstrom. *op. cit.* note 32, p. e349.

<sup>61</sup>Mehta et al., *op. cit.* note 52, p. 774; E.C. Holmes. RNA Virus Genomics: A World of Possibilities. *J Clin Invest* 2009; 119: 2488-2495.

<sup>62</sup>Hecht et al., *op. cit.* note 38, p.1240.

MSM. When HIV-1 genetic sequences originally collected for routine HIV care and treatment are used for monitoring HIV clusters, care must be taken to ensure that individual transmitters and subgroups are protected from privacy violations.<sup>63</sup> For example, Ugandan MSM face prison under their national law and the identification of a transmission cluster of MSM can provide an opportunity for criminal investigation. Furthermore, anti-MSM laws and homophobia in Africa can hamper efforts to study the spread of HIV and design of effective interventions.<sup>64</sup>

The loss of privacy and the inadvertent disclosure of HIV status for people who might have transmitted the virus in a transmission network are primary concerns underlying the sharing of HIV sequence data and transmission network analysis.<sup>65</sup> On the one hand, the uniqueness of the virus enables useful detailed molecular epidemiological analysis for clinical management (for example drug resistance testing for HIV treatment), public health interventions and prevention efforts. On the other hand, however, analysis of HIV sequence data can potentially reveal identifying information.<sup>66</sup> This is complex, because researchers have an obligation to report their findings truthfully, but also have parallel obligations to prevent and minimise foreseeable harm to participants and the community.

It is hard to quantify and preserve privacy in the context of HIV molecular epidemiology and its application in public health and management of patients. This is because of the absence of appropriate methods for assessing privacy.<sup>67</sup> Traditionally, de-identification of genetic data, which is achieved by the removal of specific identifiers, was a useful approach for the protection of privacy in research settings. However, de-identification has its own problems. Essentially, de-identification requires one to identify all the risks associated with re-identification of individuals and to establish what constitutes a threshold for safety.<sup>68</sup> This presents both conceptual and measurement challenges since quantifying the privacy properties of data is difficult in the context of genomic data.<sup>69</sup> Additionally, de-identification cannot guarantee the protection of privacy due to the ubiquitous nature of data and the increasing ability of scientists to triangulate data from different sources. De-identification also minimises the utility of the data,<sup>70</sup> particularly in molecular epidemiological studies where certain variables are required to make informed conclusions. For example, without linking HIV sequences to selected socio-demographic variables, it is not possible to generate models

that predict transmission patterns or to evaluate the impact of interventions.

### 3.6 | Independent ethics review

Prior to data collection, study protocols should be subjected to an independent, properly constituted and competent REC/IRB to give a dispassionate view of the protocol. Apart from assessing whether ethical and regulatory requirements are fulfilled, independent ethics review is also meant to check for any biases and conflicts that the researcher(s) might have. Ultimately, the review provides assurance to the public that individuals and groups will not be exploited.<sup>71</sup> The overall objective of ethics review is, therefore, to maximise protection of research participants while enhancing the quality of research. Scientific validity will also be scrutinised. The review will include appropriateness of the methods, balance between risk of harms and potential benefits and whether there are alternative and less risky methods of answering the same research question. Furthermore, the informed consent process and fair selection of study participants and how they are treated will also be examined.

Phylogenetic analysis of HIV genetic sequences is a relatively new area, which presents new ethical challenges for communities, reviewers and scientists alike, and adds new dimensions to traditional ethical concerns. These may require novel conceptual bioethics guidelines. These ethical issues have received minimal attention and relatively little has been published on the subject.<sup>72</sup> RECs may be hesitant to evaluate and provide oversight on HIV phylogenetic research because of limited expertise. They may be tempted to err on the side of caution by disapproving innovative and carefully designed HIV phylogenetic studies. Alternatively, an under-resourced REC might prematurely approve a study without requiring risk mitigation strategies. While the REC's intentions might be well meaning, such decisions potentially curtail the development of innovative research in the field. It is therefore critical to build the capacity of phylogenetic experts in research ethics and of RECs in phylogenetics so that they can make more meaningful contributions to the ethical evaluation of HIV phylogenetic studies, either as members of RECs or as independent consultants. Such approaches have also been advocated for in other types of research, for example, social science and health systems research where representation of experts in those fields on RECs has been low.<sup>73</sup> Further, in the early 2000s, the World Health Organization and UNAIDS engaged in extensive international training of RECs in anticipation of complex HIV preventive vaccine trials.<sup>74</sup>

<sup>63</sup>Gilbert M 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: e461.

<sup>64</sup>Nordling L. Homophobia and HIV Research: Under Siege. *Nature* 2014; 509: 274–275.

<sup>65</sup>Brooks, Sandstrom, *op. cit.* note 32, p. e349; Hecht et al., *op. cit.* note 38, p. 1240.

<sup>66</sup>Little et al., *op. cit.* note 13, p. e98443.

<sup>67</sup>Mehta et al., *op. cit.* note 52, p. 775.

<sup>68</sup>*Ibid.*

<sup>69</sup>Dwork C, Pottenger R. Toward Practicing Privacy. *J Am Med Inform Assoc* 2013; 20: 102–108.

<sup>70</sup>Gostin LO, Nass S. Reforming the HIPAA Privacy Rule: Safeguarding Privacy and Promoting Research. *JAMA* 2009; 301: 1373–1375.

<sup>71</sup>Emanuel et al., *op. cit.* note 20, p. 934.

<sup>72</sup>Chin WW, Wieschowski S, Prokein J, Illig T, Strech D. Ethics Reporting in Biospecimen and Genetic Research: Current Practice and Suggestions for Changes. *PLoS Biol* 2016; 14: e1002521; Brooks, Sandstrom, *op. cit.* note 32, p. e350.

<sup>73</sup>Mutenherwa, Wassenaar, *op. cit.* note 22, p.123; Wassenaar, Rattani, *op. cit.* note 23, p. 5.

<sup>74</sup>Kaleebu P et al. African AIDS Vaccine Programme for a Coordinated and Collaborative Vaccine Development Effort on the Continent. *PLoS Med* 2008; 5: e236.



### 3.7 | Informed consent

The requirement for informed consent is based on the principle of respect for persons. It ensures that research participants are given a chance to make decisions on whether they want to join a study, continue participation and whether their decision is in line with their aspirations, values, beliefs and interests. Informed consent has five main elements: information disclosure, competence, understanding, voluntariness and formalization of the agreement.<sup>75</sup>

In practical terms, informed consent requires researchers to provide prospective research participants with clear, unbiased, detailed and factual information about the study. The information disclosure ordinarily covers information on study methods, potential risks and benefits as well as assertions that participation is at the sole discretion of the research participant and that they can refuse to join or may withdraw from the study at any time without suffering negative consequences.<sup>76</sup> Information about the study risks, procedures and benefits must be disclosed to the prospective participants in a way that facilitates comprehension.

A key concern in health research globally and with particular reference to Africa is the quality of informed consent.<sup>77</sup> Empirical research has shown that participants may not receive adequate information about the study or may fail to understand research procedures and key concepts,<sup>78</sup> particularly where complex studies are conducted in resource limited settings. By their nature and design, HIV phylogenetic studies are complex, as described above. When conducted in resource-limited settings where many prospective research participants are illiterate, research-naïve and have limited access to healthcare services, such studies may be poorly understood and thus lead to invalid consent (or refusals), thereby undermining the ethical integrity of the study.

A major concern in HIV phylogenetic research is the amount of information that prospective research participants receive and how such information is best presented to guarantee valid consent.<sup>79</sup> On the one hand, detailed and inappropriately packaged information might scare prospective participants, while on the other; too little information might undermine valid consent. A related concern is that researchers have to explain complex technical terms to prospective participants. The mother tongue of some prospective participants

<sup>75</sup>Beauchamp TL, Childress JF. 2009. *Principles of biomedical ethics*. New York: Oxford University Press.

<sup>76</sup>Ibid.

<sup>77</sup>Mandava A et al. The Quality of Informed Consent: Mapping the Landscape. A Review of Empirical Data from Developing and Developed Countries. *J Med Ethics* 2012; 38: 356-365.

<sup>78</sup>Munalula-Nkandu E et al. To What did They Consent? Understanding Consent Among Low Literacy Participants in a Microbicide Feasibility Study in Mazabuka, Zambia. *Dev World Bioeth* 2014; 1: 12069; Ndebele PM, et al. Improving Understanding of Clinical Trial Procedures among Low Literacy Populations: An Intervention within a Microbicide Trial in Malawi. *BMC Med Ethics* 2012; 13: 1-14.

<sup>79</sup>Tindana P et al. Seeking Consent to Genetic and Genomic Research in a Rural Ghanaian Setting: A Qualitative Study of the MalariaGEN Experience. *BMC Med Ethics* 2012; 13: 15; Bull S, Farsides B, Tekola Ayele F. Tailoring Information Provision and Consent Processes to Research Contexts: The Value of Rapid Assessments. *J Empir Res Hum Res Ethics* 2012; 7: 37-52

may not have equivalent words for these concepts so they may be difficult to understand. Scientific terms like DNA, genes and sequencing and concepts like data sharing, use and storage of genetic material are difficult for non-experts to understand. Assessment of comprehension, as often done in complex clinical trials in developing countries<sup>80</sup> should be considered during the consent process to ascertain if adequate understanding has been achieved. Such tests are resource intensive but critical. These challenges are not new to research, but they become more prominent in HIV phylogenetics research because of its complexity.

HIV sequences used for HIV surveillance are typically generated from samples obtained for HIV diagnosis and treatment. These sequences are often used in conjunction with medical records. It must be made clear to prospective participants during the consent process what their samples and personal data will be used for. While secondary use of the data may be done for the public good, for example, HIV surveillance, prospective participants or patients should be made aware of this and of the risks and benefits involved. This is important in order to develop and maintain public trust between researchers and research participants because clinical information can potentially be abused. The fiduciary relationship between healthcare providers and patients is critical at the individual level in terms of respect for personal autonomy and reciprocity. Its absence can undermine future public health initiatives and research in the community.<sup>81</sup>

Prisoners are a vulnerable population subject to abuse and exploitation. Because of their unique circumstances, they may not be capable of giving valid informed consent to participate in health research in general and HIV phylogenetic studies in particular. Studies investigating the possible spread of HIV infection and its route of transmission among prison inmates such as the one conducted in Scotland<sup>82</sup> could potentially lead to systemic harm to prisoner subgroups. For example, injecting drug users, racial minorities and those with certain diseases and conditions that may not be treated during imprisonment could be potential targets. Adequate provisions should be put in place when conducting phylogenetics research in carceral settings to protect participants from harm.

Considering that HIV phylogenetic studies are often conducted in demographic surveillance systems and HIV treatment programmes, it is possible that prospective respondents from the community may fail to distinguish research from clinical care or developmental work. Failure to make this distinction can make voluntary participation in research problematic due to misplaced hopes and expectations - the therapeutic misconception.<sup>83</sup> Investigators

<sup>80</sup>Lindegger G et al. Beyond the Checklist: Assessing Understanding for HIV Vaccine Trial Participation in South Africa. *J Acquir Immune Defic Syndr* 2006; 43: 560-566.

<sup>81</sup>Carter P, Laurie GT, Dixon-Woods M. The Social Licence for Research: Why care.data Ran into Trouble. *J Med Ethics* 2015; 41: 404-409; Marsh V et al. Beginning Community Engagement at a Busy Biomedical Research Programme: Experiences from the KEMRI CGMRC-Wellcome Trust Research Programme, Kilifi, Kenya. *Soc Sci Med* 2008; 67: 721-733.

<sup>82</sup>Taylor A, Goldberg D, Emslie J, et al. Outbreak of HIV Infection in a Scottish prison. *BMJ* 1995; 310: 289-92.

<sup>83</sup>Appelbaum PS, Roth I, Lidz C. The Therapeutic Misconception: Informed Consent in Psychiatric Research. *Int J Law Psychiatry* 1982; 5: 319-329.

involved in HIV phylogenetic studies should make community participation an integral part of the informed consent process in order to dispel any unrealistic expectations of support from the researcher, which may influence their decision to participate in the research.

HIV sequences are generated from virions obtained from samples of HIV infected patients. These virions can be extracted from serum, plasma, semen or vaginal swabs. Once the sample leaves the participant, he or she has minimal control over what happens to it. Additionally, it is neither possible to anticipate all potential uses of the collected specimen nor predict their associated risks.

Improper use of samples is well documented.<sup>84</sup> The Arizona State University versus the Havasupai community is a classic example. In that case, community members donated their blood samples specifically for studies on diabetes, a major health problem in their community. Without consultation and obtaining valid consent, the researchers used the samples to conduct research on human genetics and population migration. In addition to conducting these unsanctioned studies, which were experienced as offensive and stigmatising, the researchers shared the samples with other researchers and published their work. Even though the samples were anonymised, the results affected the whole community, which filed a successful lawsuit accusing the researchers of fraud, breach of fiduciary duty, negligence and trespassing.

The Havasupai case illustrates that not only individuals can be harmed by research on their de-identified samples. Communities and their cultures are also vulnerable and should be protected at all costs.<sup>85</sup> More importantly, it demonstrates that boundaries of sample use should be clearly defined during the consent process to protect research participants and communities from harm.

### 3.8 | Ongoing respect for participants

Ethical research requires that the rights and welfare of research participants be respected during and after the study. Research participants cannot be separated from the communities in which they live. One cannot respect an individual and at the same time disregard their communities.

Phylogenetic research is complex as it may utilise existing databases from routine medical care, which may contain confidential social, clinical and demographic information, which could lead to the identification of individuals or social groups if not properly managed. Access and governance of such databases can pose systemic harms to individuals and communities through HIV disclosure, genetic or HIV related discrimination, stigmatisation, lawsuits and social disharmony, among others. In the absence of clear and evidence-based

guidance on access to such databases and on phylogenetic research in particular, these problems may be worsened.

HIV has been and remains stigmatised since the discovery of the virus three decades ago. Stigma and discrimination are harmful social phenomena, which affect both individuals and social groups. Those who suffer most are the already or historically marginalised, for example, women, black people, commercial sex workers, injecting drug users, sexual minorities and people from specific geographic locations. Incidents of HIV related stigma and discrimination are well documented.<sup>86</sup> Stigma and discrimination can lead to fear of HIV testing and disclosure. This fuels transmission since the infected person cannot access treatment without being tested.<sup>87</sup> In countries where HIV transmission is criminalized, knowledge of HIV status is avoided to protect individuals from prosecution. This is because if they know their HIV status, the law requires HIV positive individuals to act responsibly in order to avoid infecting others.<sup>88</sup> Failure to do so can lead to prosecution for intentionally transmitting HIV. In an effort to avoid prosecution, individuals may shy away from tests that benefit their health and the community, for example, HIV and drug resistance testing.<sup>89</sup> Stigma and discrimination are real risks in HIV transmission network analysis, as discussed above, although the problem is not unique to HIV phylogenetics.

## 4 | CONCLUSION

HIV phylogenetics is a relatively new field and limited conceptual and empirical work has been conducted to explore the ethical issues it raises. To review the body of knowledge in this field, we applied the Emanuel, Wendler and Grady Framework (EF). Its principles mimic the sequence of a research project, starting with the design and ending in the communication of the results to the participants of the study and the community at large. To the best of our knowledge, this is the first review to use the EF to analyse research in HIV phylogenetics. While the EF was instrumental in guiding the structure of this review, it was evident that some themes could fall under more than one ethical principle. For example, the discussion on sampling, which was covered under 'Fair participant selection', could equally fall under 'Scientific validity'. The same applies to issues of 'Community engagement' and 'Informed consent', which are interrelated. It is possible that other ethical considerations associated with HIV phylogenetics may not fit within the EF.

<sup>84</sup>Mello MM, Wolf LE. The Havasupai Indian Tribe case--Lessons for Research Involving Stored Biologic Samples. *N Engl J Med* 2010; 363: 204-207; Couzin-Frankel J. Ethics. Researchers to Return Blood Samples to the Yanomamo. *Science* 2010; 328: 1218; Skloot R. 2010. *The Immortal Life of Henrietta Lacks*. Crown Publishing Group.

<sup>85</sup>Dhai A, Mahomed S. Biobank Research: Time for Discussion and Debate. *S Afri Med J* 2013; 103: 225-227; Drabiak-Syed K. Lessons from Havasupai Tribe v. Arizona State University Board of Regents: Recognizing Group, Cultural, and Dignitary Harms as Legitimate Risks Warranting Integration into Research Practice. *J Health & Biomed L* 2010; VI: 175-225.

<sup>86</sup>Skinner D, Mfecane S. Stigma, Discrimination and the Implications for People Living with HIV/AIDS in South Africa. *SAHARA J* 2004; 1: 157-164.

<sup>87</sup>UNAIDS, *op. cit.* note 8, pp. 12-13; van Bogaert DK. Common Ethical Issues Related to HIV/AIDS. *S Afri Fam Pract* 2011; 53: 605-609.

<sup>88</sup>Weait M. 2007. *Intimacy and Responsibility: The Criminalisation of HIV Transmission*. London: Routledge; J. Chalmers. 2013. *Disease Transmission, Liability and Criminal Law*. In: *Criminal Law, Philosophy and Public Health Practice*. Viens AM, Coggon J, Kessel AS, editors. Cambridge: Cambridge University Press.

<sup>89</sup>Brooks, Sandstrom, *op. cit.* note 32, p. e350.

We recommend that this review be followed by an empirical investigation of stakeholders' perspectives (that is participants, community, researchers and ethicists) on key ethical issues on HIV phylogenetics. In spite of the increasing number of large HIV phylogenetic studies in the developed and developing world, clear guidelines are still not available. Empirical work is needed to identify best practices regarding how researchers and public health practitioners convey key messages about HIV phylogenetics. This needs to be done with a degree of simplicity and accuracy so that prospective participants can understand the research protocol and make informed decisions so that valid consent can be obtained. It is also important to develop guidelines that support the rapid sharing of data so that effective interventions can be produced to halt the spread of HIV outbreaks. The rapid availability of sequence data during the Ebola epidemic was one of the factors that helped to control the explosive outbreak in West Africa.<sup>90</sup> However, the guidelines also need to safeguard the privacy of HIV infected individuals, and possibly of host communities. Such guidelines require novel conceptual and empirical work to supplement existing research ethics scholarship and resources. Best practices for community engagement<sup>91</sup> also need to be developed and evaluated for large community-based phylogenetic projects. In conclusion, we believe that novel conceptual and empirical ethics research needs to be conducted in order to complement the initial analysis presented in this paper and to increase the favourable risk-benefit ratio of HIV phylogenetic studies.

## FUNDING

The authors would like to acknowledge financial support from the Wellcome Trust Strategic Award (096527).

## CONFLICT OF INTEREST

No conflicts declared.

## ORCID

Farirai Mutenherwa  <http://orcid.org/0000-0001-7443-7676>

**Farirai Mutenherwa** is a PhD student in the School of Applied Human Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa, and the KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa.

**Douglas Wassenaar** is Professor and Director, South African Research Ethics Training Initiative (SARETI) and a consultant to the HIV AIDS Vaccines Ethics Group (HAVEG) in the School of Applied Human Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa.

**Tulio de Oliveira** is Professor in the School of Laboratory Medicine and Medical Sciences, Nelson Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, Research Associate at the Centre for the AIDS Programme of Research in South Africa (CAPRISA) & Director of KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

**How to cite this article:** 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. *Developing World Bioeth.* 2018;00:1-11. <https://doi.org/10.1111/dewb.12191>

<sup>90</sup>Quick J et al. Real-time, Portable Genome Sequencing for Ebola Surveillance. *Nature* 2016; 530: 228-232.

<sup>91</sup>UNAIDS-AVAC, *op. cit.* note 26, p. 5-63.