



# BMJ Open Beyond HIV outbreaks: protocol, rationale and implementation of a prospective study quantifying the benefit of incorporating viral sequence clustering analysis into routine public health interventions

Jon A Steingrimsson ,<sup>1</sup> John Fulton,<sup>2</sup> Mark Howison,<sup>3</sup> Vlad Novitsky,<sup>4</sup> Fizza S Gillani,<sup>4</sup> Thomas Bertrand,<sup>5</sup> Anna Civitarese,<sup>5</sup> Katharine Howe,<sup>5</sup> Guillermo Ronquillo,<sup>5</sup> Benjamin Lafazia,<sup>5</sup> Zoanne Parillo,<sup>5</sup> Theodore Marak,<sup>5</sup> Philip A Chan,<sup>4,5</sup> Lila Bhattarai,<sup>5</sup> Casey Dunn,<sup>6</sup> Utpala Bandy,<sup>5</sup> Nicole Alexander Scott,<sup>5</sup> Joseph W Hogan,<sup>1</sup> Rami Kantor <sup>4</sup>

**To cite:** Steingrimsson JA, Fulton J, Howison M, *et al*. Beyond HIV outbreaks: protocol, rationale and implementation of a prospective study quantifying the benefit of incorporating viral sequence clustering analysis into routine public health interventions. *BMJ Open* 2022;**12**:e060184. doi:10.1136/bmjopen-2021-060184

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-060184>).

Received 16 December 2021  
Accepted 29 March 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Rami Kantor;  
[rkantor@brown.edu](mailto:rkantor@brown.edu)

## ABSTRACT

**Introduction** HIV continues to have great impact on millions of lives. Novel methods are needed to disrupt HIV transmission networks. In the USA, public health departments routinely conduct contact tracing and partner services and interview newly HIV-diagnosed index cases to obtain information on social networks and guide prevention interventions. Sequence clustering methods able to infer HIV networks have been used to investigate and halt outbreaks. Incorporation of such methods into routine, not only outbreak-driven, contact tracing and partner services holds promise for further disruption of HIV transmissions.

**Methods and analysis** Building on a strong academic–public health collaboration in Rhode Island, we designed and have implemented a state-wide prospective study to evaluate an intervention that incorporates real-time HIV molecular clustering information with routine contact tracing and partner services. We present the rationale and study design of our approach to integrate sequence clustering methods into routine public health interventions as well as related important ethical considerations. This prospective study addresses key questions about the benefit of incorporating a clustering analysis triggered intervention into the routine workflow of public health departments, going beyond outbreak-only circumstances. By developing an intervention triggered by, and incorporating information from, viral sequence clustering analysis, and evaluating it with a novel design that avoids randomisation while allowing for methods comparison, we are confident that this study will inform how viral sequence clustering analysis can be routinely integrated into public health to support the ending of the HIV pandemic in the USA and beyond.

**Ethics and dissemination** The study was approved by both the Lifespan and Rhode Island Department of Health Human Subjects Research Institutional Review Boards and study results will be published in peer-reviewed journals.

## Strengths and limitations of this study

- The first study to design, implement, and evaluate a system that incorporates real-time HIV molecular clustering information into routine contact tracing and partner service efforts within a health department.
- The study will be able to draw causal conclusions about the impact of different clustering methods without randomising index cases to clustering methods, enabling optimal use of information.
- The study has a four-step ethics plan to ensure that all procedures conform to ethical standards.
- The study is done in a single state preventing cross-jurisdictional transmission networks to be incorporated.

## INTRODUCTION

In 2019, thirty-eight million people worldwide were living with HIV,<sup>1</sup> 1.7 million of whom were newly infected and about 700 000 died from AIDS-related illnesses. In the USA, approximately 1.2 million are currently living with HIV, with disproportionate impact on racial/ethnic minorities and gay and bisexual men.<sup>2</sup> Although rates of new infections are decreasing in recent years, infection numbers have remained stable and the Centers for Disease Control and Prevention (CDC) estimates that 14% of those living with HIV are unaware of their infection. The United Nations set the ambitious goal to end the AIDS pandemic by 2030,<sup>3</sup> which the USA has adopted.<sup>4</sup> To achieve that goal, novel methods to disrupt HIV transmission are needed.

Phylogenetics, the study of evolutionary relationships between organisms,<sup>5</sup> can be used to identify grouping of individual-level HIV sequences termed clusters,<sup>6</sup> which may provide (non-causal) information about transmission networks.<sup>7–11</sup> These methods can be applied to identify HIV clusters using sequence data routinely collected for HIV drug resistance testing.

Near real-time clustering analyses have been successfully used to identify outbreaks and guide public health responses to outbreak investigations.<sup>8,9,12,13</sup> For example, in 2015, an investigation of an HIV outbreak was initiated by a disease intervention specialist (DIS) that identified a cluster of 11 new HIV diagnoses in Scott County, Indiana. Phylogenetic analysis of HIV sequences revealed a cluster of 181 cases, the majority of whom (88%) reported having used oxymorphone intravenously. Contact tracing and partner service efforts resulted in 536 named partners of whom 468 were located, tested for HIV and, if HIV positive, linked to care. As a result, Indiana established a syringe service programme for the first time.<sup>12</sup>

Rapid response to emerging clusters was one of the four focus points identified by the US Department of Health and Human Services to disrupt HIV transmission.<sup>4</sup> To assist with cluster identification, the CDC developed, manages, and distributes HIV-TRACE, a tool for identifying transmission clusters, currently incorporated into CDC guidelines for outbreak investigations.<sup>14</sup>

Despite such advancements, and as we previously introduced,<sup>15</sup> several key questions remain in the application of these methods to public health problems. First, how to optimally respond to information derived from sequence clustering methods to improve the yield of contact tracing and partner notification and disrupt HIV transmission? Second, can and should such methods be used in routine, rather than only outbreak-driven, workflows of public health departments? And third, does effectiveness in identifying transmission networks vary between different clustering methods? As it is unlikely that any data can reveal the actual transmission network, it is important to develop and implement studies that address these questions.

Up to now, studies incorporating clustering methods have been mostly conducted in research environments, primarily to address outbreak scenarios or characterize existing HIV epidemics. This study was instituted to design, implement and evaluate a system to answer these questions in normal, day-to-day public health contact tracing and partner service practices, that is, beyond the limits of traditional single outbreak investigations. We provide an outline and a rationale for the study; develop a cluster analysis triggered intervention (CATI) to be used in routine public health activities; plan the evaluation of this intervention and of its effects on prioritising the deployment of contact tracing and partner service resources; sketch a novel study design that facilitates the evaluation of alternative clustering methods while avoiding randomisation to a single method, which enables the most efficient use of information and is important in

a small state like Rhode Island (RI); and address ethical questions which naturally arise when such data are used to guide public health interventions.

## RATIONALE AND STUDY AIMS

A major challenge to disrupting HIV transmission is that actual transmission networks are unknown and can only be inferred loosely from observed data. Greater precision of inference can be obtained by combining information from two sources: standard, public health contact tracing and partner services, and cluster analyses of HIV sequences. The first provides information on social and sexual contacts (social networks)<sup>16</sup> and the second provides information about virus similarity and evolution<sup>17,18</sup> from which additional inferences about social networks may be hypothesised.

Contact tracing and partner services is the action of interviewing newly HIV-diagnosed index cases (ICs) to identify relationships in social networks.<sup>19</sup> Any partner of an IC newly identified through contact tracing and partner services is notified, offered HIV testing and, if positive, linked to care. Currently, contact tracing and partner services is a standard public health tool for preventing HIV transmission, but it is limited in that ICs and their partners may either be unwilling to provide partner information or may be unable to do so (eg, some contacts may be anonymous while others may simply be forgotten in the course of time). Furthermore, many public health agencies are so limited in resources that they cannot perform contact tracing and partner services on all newly identified ICs, thereby foregoing much information about the scope and nature of the social networks in which HIV is transmitted.<sup>20</sup>

We hypothesise that routinely conducting cluster analyses of HIV sequences in a state-wide epidemic provides important information about potential transmission networks; that such information can be used to develop an intervention for more comprehensive and/or targeted tracing and partner services; that these improved interventions might yield uninfected individuals at high risk of getting infected, individuals who are infected but unaware of their infection, and ‘infected-aware’ individuals who are not linked to care, that would not be identified by routine contact tracing and partner service efforts; and that the choice of clustering method to identify clusters may be important.

To address these hypotheses, we designed this study, in which we will (1) conduct real-time and routine clustering analyses to detect ICs that cluster in a state-wide HIV epidemic; (2) design and implement a public health intervention triggered by the integration of this real-time sequence clustering information with routine, rather than outbreak-driven, contact tracing and partner services; and (3) quantify the benefit of the implementation of this real-time public health intervention. By executing this study, we will determine the yield of real-time CATI; evaluate how different methods used to identify clusters

impact that yield; identify strategies or factors that are associated with higher yields of CATI; and identify ways to direct contact tracing and partner service resources towards individuals who are likely to yield high returns in terms of disruption of HIV transmission.

The goal of integrating cluster analysis information with routine contact tracing and partner services also raises ethical questions<sup>21-23</sup> that must be considered. Therefore, we have developed a four-step ethics platform as an integral part of the study to ensure that all procedures we implement or alter conform to ethical standards.

### Study setting

The study is conducted in RI, a state with population of 1.1 million, where the RI Department of Health (RIDOH) has for years conducted routine contact tracing and partner services among *all* new ICs. RI has achieved significant success addressing HIV over the last two decades,<sup>24</sup> but the number of new infections in the state has stabilized,<sup>25 26</sup> primarily because of transmission in high-risk networks, particularly among men who have sex with men (MSM).

Our study includes all adults  $\geq 18$  years old, who are newly HIV diagnosed after 1 January 2021 and receive HIV care at an RI healthcare facility. This study will involve all eligible ICs over a 2-year period. Based on data from prior years we expect a sample size of around 180 ICs. Collaborators in the effort include RIDOH, the Miriam Hospital Immunology Center (the largest RI HIV center, caring for ~80% of the state's HIV population), Brown University School of Public Health, and expert consultants.

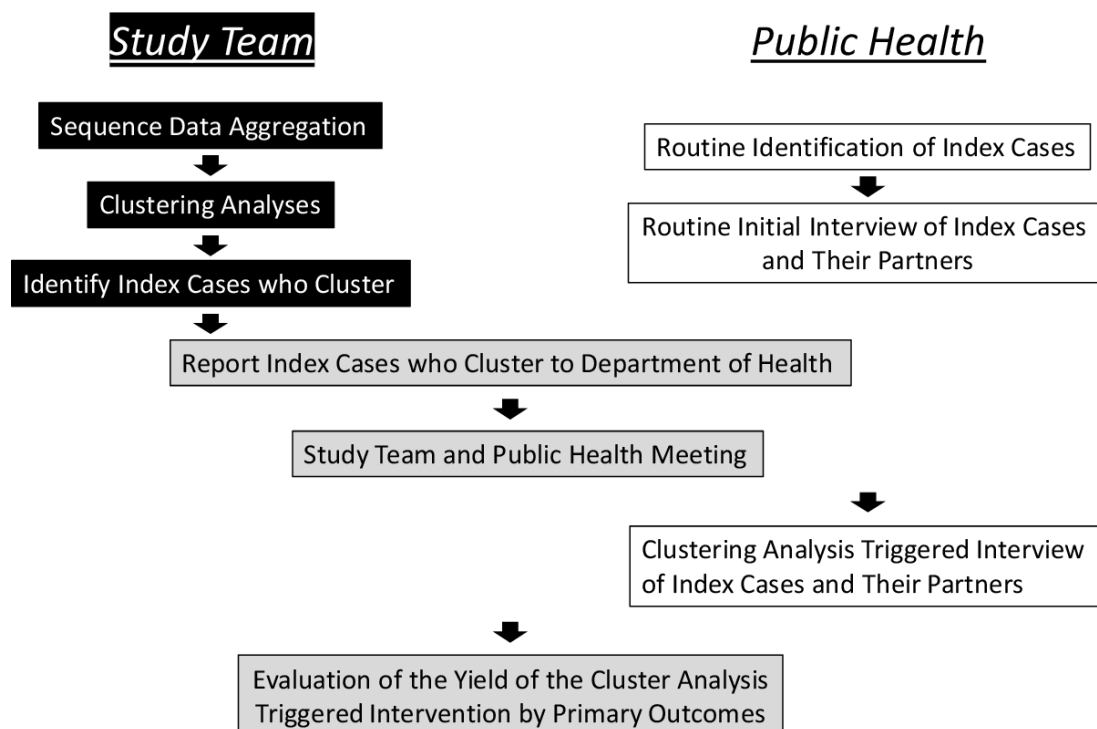
### Study overview

We longitudinally collect state-wide HIV sequence data available from routine drug resistance testing in clinical care; conduct real-time clustering analyses with multiple methods to infer presence of ICs in clusters; have real-time discussions of results among academic and public health collaborators; implement in real time a CATI; and evaluate the benefit of this integration and intervention. **Figure 1** provides schematic representation of this process, and below, we provide detailed descriptions of study aims and the ethics platform that encompasses them.

### Conduct real-time and routine sequence clustering analyses to detect ICs that cluster in a state-wide HIV epidemic

To enable real-time integration of information derived from HIV sequence cluster analysis into routine contact tracing and partner services, we will develop an automated pipeline that, on a monthly basis, aggregates HIV-1 state-wide sequence data, performs sequence quality control, identifies clusters, links sequence and clinical data, identifies ICs that cluster and reports them to RIDOH, to be followed by CATI.

State-wide sequencing of the partial HIV-1 *pol* gene routinely performed for drug resistance testing in those receiving care at an RI facility will be aggregated, as has been done since 2003, including people HIV diagnosed as early as the 1980s.<sup>17 27-30</sup> This process will include all historically collected as well as all prospectively available, newly generated, monthly sequence data.



**Figure 1** Schematic representation of the monthly study process. Black boxes represent processes that involve the study team, white boxes represent processes that involve the Rhode Island Department of Health, and grey boxes represent processes that involve both the study team and the Rhode Island Department of Health.

Each month, when a new sequence dataset becomes available, multiple clustering approaches will be used to identify clusters. Approaches will include an ensemble of commonly used phylogenetic methods and cluster-defining characteristics and HIV-TRACE.<sup>31</sup> Methods and thresholds were selected based on an empirical comparison of approaches for identifying molecular HIV clusters.<sup>32</sup>

Once cluster analyses are complete, a list of all cluster-associated ICs (CAIC) for the month will be generated, where an IC is considered a CAIC, if it clusters with at least one method in the ensemble or HIV-TRACE. The pipeline currently takes approximately 24–72 hours to run on a computer cluster using up to 100 central processing unit cores.

### Patient and public involvement

No patients were involved in the design or conduct of the study.

#### Design and implement a public health intervention triggered by the integration of real-time sequence clustering information with routine, rather than outbreak-driven, contact tracing and partner services

To incorporate information from sequence clustering into routine contact tracing and partner services, a specific CATI has been designed and will be implemented for all CAICs. The approach has five steps:

1. RIDOH conducts contact tracing and partner services for each IC newly HIV diagnosed in RI. In this routinely performed ‘initial interview’, a DIS interviews ICs to identify at-risk contacts like sexual or needle-sharing partners. Named at-risk contacts who can be located are invited for consultation, offered HIV testing and linked to care if HIV positive but disengaged.
2. Each month, once new state-wide HIV sequence data are available and analyses are completed, the study team will provide RIDOH approved personnel with the list of CAICs using secure password-protected Web File Repository tools, as routinely done in other scenarios. RIDOH personnel will be blinded to the methods by which the ICs clustered.
3. Each month, after step 2, the study team and RIDOH will meet to discuss new CAICs, and review past CAICs previously discussed. Discussions will include clustering, demographic, clinical, and laboratory information, all fully deidentified.
4. Within 7 days of the monthly meeting, RIDOH’s DIS will contact all new CAICs to invite them to participate in the CATI, a reinterview, whose goals are to elicit additional information on the CAIC’s recent sexual and needle-sharing partners, while integrating the new cluster analysis information. If the DIS locates the CAIC, a face-to-face or a phone meeting will be attempted. During the reinterview, the DIS will notify the CAIC that since they last spoke/met they learnt that the HIV strain of the CAIC has become increasingly common in RI, and that when this happens, they reach back out

to people to share these data, confirm what they have shared earlier and re-elicite sexual and needle-sharing partners.

5. The DIS will attempt to contact each newly identified partner obtained from the CAIC in the CATI, using routine sources like LexisNexis<sup>33</sup> and social media to facilitate searches. If the partner is located, a face-to-face meeting will be attempted. If the partner is not an RI resident, the DIS will forward their contact information to the appropriate state health department for follow-up. During the partner interview, the DIS will notify them of the recent HIV exposure and that the HIV strain they may have been exposed to has become increasingly common in RI, and that they will work with them to ensure they get tested. If HIV diagnosis is confirmed, a new investigation begins, with the partner as an IC.

ICs and their named partners will never be informed about who named them, a standard procedure for contact tracing and partner services,<sup>34</sup> and no information will be provided to the IC or their named partners on other cluster members.

#### Quantify the benefit of implementation of real-time and routine CATI to disrupt HIV transmission in RI

The study is designed to enable evaluation of the benefit of real-time integration of information from cluster analysis and routine contact tracing and partner services to public health. We will compare the following contact tracing and partner service approaches:

1. Routine contact tracing and partner services with no CATI (no reinterviews).
2. Reinterviewing ICs that cluster by using HIV-TRACE.
3. Reinterviewing ICs that cluster by using a phylogenetic ensemble.

The primary endpoints used to quantify differences between the three approaches are:

1. Number of CAICs naming more than one sexual or needle-sharing partner.
2. Number of notifiable partners elicited from CAICs.
3. Number of notifiable partners of ICs with HIV, who have become ICs.
4. Number of partners of ICs whose HIV status is newly determined.

For all evaluation measures, the yield of routine (not cluster analysis triggered) contact tracing and partner services is the yield of the initial interview; the yield of phylogenetic CATI is the combined yield of the initial interview and the reinterview (if the IC clusters with the phylogenetic ensemble); and the yield of HIV-TRACE CATI is the combined yield of the initial interview and the reinterview (if the IC clusters with HIV-TRACE).

Throughout the CATI process, data will be collected from health records and databases to allow determination of the outcomes and comparison to the initial interview yield. Additional secondary objectives will include comparing the yield of individual components of the ensemble and identifying factors associated with

higher CATI yields. The latter can help direct public health resources towards individuals likely to yield high returns in disrupting HIV transmission. Previously,<sup>32</sup> we conducted empirical comparisons of approaches for identifying clusters. Due to lack of data on the impact of integrating phylogenetic clustering methods into public health efforts, the empirical comparison focused on concordance between different approaches rather than the more relevant public health question of what methods are most effective at disrupting transmissions, a question that this study can answer.

### Statistical analysis

A key study design component is combining clustering methods together and identifying CAICs by any method, and then use combined results to inform the CATI. By doing this, we can ascertain what downstream actions would have been taken if only one method was used—that is, we can attribute contacts and associated actions to the method(s) that identified them. This permits causal comparisons between different clustering methods without randomising ICs to a clustering method, and without interfering with DIS work. As a result, all ICs contribute data to all clustering methods making the study feasible with a smaller sample size.

The comparisons between the three contact tracing and partner service approaches will be tested using a generalised linear mixed effects model with the evaluation measure as the outcome, the method ((a) routine contact tracing and partner services, or (b) contact tracing and partner services that incorporate HIV-TRACE, or (c) contact tracing and partner services that incorporate phylogenetic ensemble) as a fixed effect and the IC as a random effect. Two-sided hypothesis tests will be used when comparing the phylogenetic ensemble and HIV-TRACE. As by design, HIV-TRACE and the phylogenetic ensemble cannot have a lower yield than routine contact tracing and partner services, all comparisons to routine contact tracing and partner services will be one sided.

If we assume that 40% of ICs name  $\geq 1$  partner during initial interview (based on historical data), that 45% of ICs name  $\geq 1$  partner during initial interview or HIV-TRACE-initiated reinterview, and that 50% of ICs name  $\geq 1$  partner during initial interview or phylogenetic ensemble-initiated reinterview, then with 180 ICs we will have  $>80\%$  power to test all three hypotheses.

If we assume that the effect size comparing the average number of notifiable partners elicited from ICs is  $\geq 0.25$  for all three comparisons, then with 180 ICs we will have  $>80\%$  power to test all three hypotheses (variances estimated using historical data). If we assume that the number of HIV tests per IC from routine contact tracing is 0.34 (from historical data) and that the effect size for all comparisons is  $\geq 0.17$ , then with 180 CAICs we will have  $>80\%$  power for testing all three hypotheses. A study design that randomises 180 ICs to one of the three partner service approaches has  $<45\%$  power for testing all three hypotheses.

### Ethics statement

The study was approved by the Lifespan–Miriam Hospital Institutional Review Board, Providence, RI (ID FWA00003538) and the RIDOH Institutional Review Board, Providence, RI (ID FWA00006141).

### ETHICS AND DISSEMINATION

We provide this brief section to acknowledge the importance of ethics consideration within this topic and our special attention to it as we conduct our study. To assure that our study methodology meets strict ethical standards, we followed a four-step agenda in building an ‘ethics platform’ on which our methods related to contact tracing and partner services are founded, and by means of which they are tested.

In step 1, we inventoried state and federal laws and related regulations which govern the elicitation and use of information from contact tracing and partner services and described our proposed methodology to a diverse group of community stakeholders and invited their comments.

In step 2, we refined our methodology to assure its conformance with comments from step 1 and submitted a detailed description of it to the two institutional review boards (IRBs) under whose aegis our study is being conducted. Our methods were further refined in discussions with these bodies and approved by them.

In step 3, we adopted an IRB-approved business associate agreement between RIDOH and the principal investigator (RK), in which the latter became an agent of the former in the management and analysis of all HIV surveillance information collected by RIDOH.

In step 4, we established an ongoing ethics review of all aspects of the study, focusing on *respect for subjects* and the *protection of personally identifiable data*. With regard to the former, we created a form with which subject *engagement* in first and second contact tracing and partner service interviews is evaluated by RIDOH staff, with the notion that scrupulous attention to *respect for subjects* would be evidenced by positive engagement of subjects during interviews. With regard to the latter, we review the conformity of all data transfer, management, and analysis to security protocols on a monthly basis.

The ethical issues contained in our study are well known and commonly managed in public health settings. However, as we develop a comprehensive integration of data from sequence clustering analysis and contact tracing and partner service interviews, we will ensure that our ethics platform grows and develops through multiple cycles of steps 1 through 4, as our research agenda develops on the basis of results. On completion of the study, the results will be published in a peer-reviewed journal.

### DISCUSSION

We provide an outline and a rationale for a study to develop methods for, and execute evaluation of, a real-time public health intervention triggered by integration



of information from sequence clustering analysis into routine contact tracing and partner services, with special attention to associated ethical aspects. With a strong academic–public health collaboration, and using routinely collected viral sequence data and routinely performed public health activities, we hypothesize that this approach will be beneficial to disrupt HIV transmissions in RI.

A major strength of the approach presented here includes the incorporation of information from sequence clustering methods in routine, rather than only outbreak-driven public health activities, under the hypothesis that doing so will further disrupt HIV transmissions. Other major strengths include the small size of RI, the state-wide representation of its HIV epidemic, the availability of longitudinal HIV sequence data and the strong and close academic–public health relationship and collaboration.

There are also several potential limitations to our study design and intervention. First, the study is only being done in a single state, and region specific and cross-jurisdictional transmission networks cannot be incorporated. Therefore, despite the in-depth state-wide implementation, generalizability of results may be limited. Second, the small size of RI and the fact that RIDOH has already been conducting routine contact tracing and partner services on all ICs may negatively impact the pace of the study and the yield expected from CATI. Third, inevitably and as we previously recognized,<sup>15</sup> the study is being done in the context of other public health prevention activities, such as educational activities and condom promotions, making isolation of the CATI effect challenging. Fourth, the CATI relies on access to sequencing data, which unfortunately at the present time are not routinely available in many resource-limited settings. Lastly, we are only conducting reinterviews for CAICs; no reinterviews are conducted for ICs that do not cluster. This prevents us from comparing the yield of cluster-based reinterviews to a control group that is randomly selected for reinterviews.

In summary, the study presented here can address key unanswered questions about the benefit of integration of a CATI into routine, rather than only outbreak-driven contact tracing and partner services. We developed an intervention (reinterviewing ICs that cluster) triggered by clustering analysis which incorporates clustering information into the interview itself, and evaluate the intervention with a novel design that avoids randomisation and allows method comparison. Whether and how this approach is beneficial, as well as whether the small size of RI is an advantage or a disadvantage in this setting, remains to be determined. Importantly, we are confident that results from this study will inform public health contact-tracing practices and add to our understanding of how information from sequence clustering analysis can be integrated into HIV care and public health interventions, supporting the goal of ending the HIV pandemic in the USA and beyond.

#### Author affiliations

- <sup>1</sup>BioStatistics, Brown University School of Public Health, Providence, Rhode Island, USA  
<sup>2</sup>Department of Behavioral and Social Sciences, Brown University, Providence, Rhode Island, USA  
<sup>3</sup>Research Improving People's Lives, Providence, Rhode Island, USA  
<sup>4</sup>Department of Medicine, Brown University, Providence, Rhode Island, USA  
<sup>5</sup>Rhode Island Department of Health, Providence, Rhode Island, USA  
<sup>6</sup>Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, USA

**Contributors** JAS and RK conceived the study. JAS, RK and JF wrote the first draft. RK, JAS, JWH, VN, MH, CD and FSG provided statistical and computational expertise. JF provided ethical content expertise. JAS, JF, MH, VN, FSG, TB, AC, KH, GR, BL, ZP, TM, PAC, LB, CD, UB, NAS, JWH and RK provided feedback on the study protocol, intellectual content and feasibility of implementation, and read and approved the final version.

**Funding** The study was partly funded by NIH grants R01AI136058, K24AI134359 and P30AI042853, and partially supported by computational resources and services at the Center for Computation and Visualization, Brown University.

**Competing interests** MH reports fees from Competition Economics and The Miriam Hospital for consulting, outside the submitted work. RK reports receiving a research grant from Gilead Sciences for work that is not related to HIV or this paper.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Jon A Steingrimsson <http://orcid.org/0000-0003-2116-9377>  
 Rami Kantor <http://orcid.org/0000-0003-3659-1108>

#### REFERENCES

- 1 Uniaids. Preliminary UNAIDS 2021 epidemiological estimates, 2021. Available: [https://www.unaids.org/sites/default/files/media\\_asset/UNAIDS\\_FactSheet\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf)
- 2 Centers for disease control and prevention. estimated HIV incidence and prevalence in the United States, 2010–2015. HIV surveillance supplemental report, 2020. Available: <https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-25-1.pdf> [Accessed 16 Nov 2021].
- 3 United Nations General Assembly. Political Declaration on HIV and AIDS: on the fast-track to accelerate the fight against HIV and to end the AIDS epidemic by 2030, 2016. Available: [https://www.unaids.org/sites/default/files/media\\_asset/2016-political-declaration-HIV-AIDS\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/2016-political-declaration-HIV-AIDS_en.pdf)
- 4 Fauci AS, Redfield RR, Sigounas G. Ending the HIV epidemic: a plan for the United States. *Jama* 2019;321.
- 5 Nei M, Kumar S. *Molecular evolution and phylogenetics*. Oxford university press, 2000.
- 6 Castro-Nallar E, Pérez-Losada M, Burton GF, *et al*. The evolution of HIV: inferences using phylogenetics. *Mol Phylogenet Evol* 2012;62:777–92.
- 7 Oster AM, France AM, Panneer N, *et al*. Identifying clusters of recent and rapid HIV transmission through analysis of molecular surveillance data. *J Acquir Immune Defic Syndr* 2018;79:543–50.
- 8 Poon AFY, Gustafson R, Daly P, *et al*. Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: an implementation case study. *Lancet HIV* 2016;3:e231–8.
- 9 Little SJ, Kosakovsky Pond SL, Anderson CM, *et al*. Using HIV networks to inform real time prevention interventions. *PLoS One* 2014;9:e98443.
- 10 Wertheim JO, Oster AM, Hernandez AL, *et al*. The International dimension of the U.S. HIV transmission network and onward

- transmission of HIV recently imported into the United States. *AIDS Res Hum Retroviruses* 2016;32:1046–53.
- 11 Aldous JL, Pond SK, Poon A, *et al*. Characterizing HIV transmission networks across the United States. *Clin Infect Dis* 2012;55:1135–43.
  - 12 Peters PJ, Pontones P, Hoover KW, *et al*. Hiv infection linked to injection use of oxymorphone in Indiana, 2014–2015. *N Engl J Med* 2016;375:229–39.
  - 13 Brenner BG, Roger M, Stephens D, *et al*. Transmission clustering drives the onward spread of the HIV epidemic among men who have sex with men in Quebec. *J Infect Dis* 2011;204:1115–9.
  - 14 Centers for Disease Control and Prevention. Detecting and responding to HIV transmission clusters: a guide for health departments. draft version, 2018. Available: <https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/cdc-hiv-ps18-1802-attachemente-detecting-investigating-and-responding-to-hiv-transmission-clusters.pdf> [Accessed 16 Nov 2021].
  - 15 Kantor R, Fulton JP, Steingrimsson J, *et al*. Challenges in evaluating the use of viral sequence data to identify HIV transmission networks for public health. *Stat Commun Infect Dis* 2020;12. doi:10.1515/scid-2019-0019. [Epub ahead of print: 11 11 2020].
  - 16 Klovdahl AS. Social networks and the spread of infectious diseases: the AIDS example. *Soc Sci Med* 1985;21:1203–16.
  - 17 Chan PA, Hogan JW, Huang A, *et al*. Phylogenetic investigation of a statewide HIV-1 epidemic reveals ongoing and active transmission networks among men who have sex with men. *J Acquir Immune Defic Syndr* 2015;70:428–35.
  - 18 Grabowski MK, Herbeck JT, Poon AFY. Genetic cluster analysis for HIV prevention. *Curr HIV/AIDS Rep* 2018;15:182–9.
  - 19 Hogben M, Niccolai LM. Innovations in sexually transmitted disease partner services. *Curr Infect Dis Rep* 2009;11:148–54.
  - 20 Armbruster B, Brandeau ML. Contact tracing to control infectious disease: when enough is enough. *Health Care Manag Sci* 2007;10:341–55.
  - 21 Mehta SR, Schairer C, Little S. Ethical issues in HIV phylogenetics and molecular epidemiology. *Curr Opin HIV AIDS* 2019;14:221–6.
  - 22 Wertheim JO, Morris S, Ragonnet-Cronin M. Consent and criminalisation concerns over phylogenetic analysis of surveillance data - Authors' reply. *Lancet HIV* 2019;6:e420–1.
  - 23 Romero-Severson E, Nasir A, Leitner T. What should health departments do with HIV sequence data? *Viruses* 2020;12:1018.
  - 24 Chan PA, Montgomery M, Marak T, *et al*. A Nearly 50% Decrease in New HIV Diagnoses in Rhode Island from 2006–2016: Implications for Policy Development and Prevention. *R I Med J* 2018;101:41.
  - 25 Rhode island department of health. HIV: getting to zero in Rhode island, 2013. Available: <https://health.ri.gov/publications/epidemiologicalprofiles/2013HIVAIDSViralHepatitisWithSurrogateData.pdf>
  - 26 Rhode Island Department of Health. 2013 Rhode island HIV/AIDS/ Viral hepatitis epidemiologic profile with surrogate data, 2013. Available: <https://health.ri.gov/publications/epidemiologicalprofiles/2013HIVAIDSViralHepatitisWithSurrogateData.pdf>
  - 27 Rogo T, DeLong AK, Chan P, *et al*. Antiretroviral treatment failure, drug resistance, and subtype diversity in the only pediatric HIV clinic in Rhode island. *Clin Infect Dis* 2015;60:1426–35.
  - 28 Chan PA, Reitsma MB, DeLong A, *et al*. Phylogenetic and geospatial evaluation of HIV-1 subtype diversity at the largest HIV center in Rhode island. *Infect Genet Evol* 2014;28:358–66.
  - 29 Chan PA, Kazi S, Rana A, *et al*. Short communication: new HIV infections at southern new England academic institutions: implications for prevention. *AIDS Res Hum Retroviruses* 2013;29:25–9.
  - 30 Chan PA, Tashima K, Cartwright CP, *et al*. Short communication: transmitted drug resistance and molecular epidemiology in antiretroviral naive HIV type 1-infected patients in Rhode island. *AIDS Res Hum Retroviruses* 2011;27:275–81.
  - 31 Kosakovsky Pond SL, Weaver S, Leigh Brown AJ, *et al*. HIV-TRACE (transmission cluster engine): a tool for large scale molecular epidemiology of HIV-1 and other rapidly evolving pathogens. *Mol Biol Evol* 2018;35:1812–9.
  - 32 Novitsky V, Steingrimsson JA, Howison M, *et al*. Empirical comparison of analytical approaches for identifying molecular HIV-1 clusters. *Sci Rep* 2020;10:1–11.
  - 33 Weaver DA, Bimber B. Finding news stories: a comparison of searches using LexisNexis and Google news. *Journal Mass Commun Q* 2008;85:515–30.
  - 34 Centers for Disease Control and Prevention. *Recommendations for partner services programs for HIV infection, syphilis, gonorrhea and chlamydial infection*, 2008.