Demographic characteristics of sources of HIV-1 transmission in Zambia

Matthew Hall[1], Tanya Golubchik[1], David Bonsall[1,2], Lucie Abeler-Dörner[1], Mohammed Limbada[3,4], Barry Kosloff[3,4], Ab Schaap[3], Mariateresa de Cesare[1,2], George Mackintyre-Cockett[1,2], William Probert[1], Oliver Ratmann[5], Ana Bulas Cruz[1], Estelle Piwowar-Manning[6], David N Burns[7], Myron S Cohen[8], Deborah J Donnell[9], Sue H Eshleman[6], Musonda Simwinga[3], Richard Hayes[4], Sarah Fidler[10], Helen Ayles[3,4], Christophe Fraser[1] on behalf of the HPTN 071 (PopART) Phylogenetics protocol team and the PANGEA protocol team

[1] Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[2] Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[3] Zambart, University of Zambia, Lusaka, Zambia
[4] London School of Hygiene and Tropical Medicine, London, United Kingdom
[5] Department of Mathematics, Imperial College, London, UK
[6] Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA
[7] Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, Maryland, United States of America.
[8] Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, USA
[9] Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[10] Department of Infectious Disease Epidemiology, Imperial College, London, UK

Abstract

BACKGROUND

In the last decade, universally available antiretroviral therapy has led to reduced HIV incidence in sub-Saharan Africa. Sources of remaining transmission need to be characterised to design effective prevention strategies.

METHODS

We used phylogenetics to understand the population characteristics of people who are sources of infection. HIV samples from 6,864 individuals from Zambia were deep-sequenced as part of HPTN 071-02 (PopART) Phylogenetics between 2014 and 2018. We identified 300 likely directed transmission pairs and analysed their sources to better understand transmission in the general population.

RESULTS

After demographic weighting of the recipient population to match the estimated total population infected during the trial period, 59.4% (95% CI: 53.1%-65.8%) of transmissions were male-to-female, with 43.1% (36.6%-49.5%) of transmissions from males aged 25-40. Since the adult HIV prevalence was 2.0 times higher in women than men, the per-capita transmission rate was 2.93 times higher per infected male than per infected female. 25.6% (19.9%-31.3%) of sources were estimated to have themselves been infected less than a year before the transmission event. 16.8% (12%-21.7%) of sources transmitted viruses resistant to first-line ART. 13% (8.6%-17.4%) of transmissions occurred between individuals from different study communities.

CONCLUSIONS

Our findings suggest that HIV transmission in the study communities took place as part of common sexual mixing, and that there was no outsized contribution of importation, from drug resistance, or recent infection. Men aged 25-40 should be the focus of treatment for prevention interventions, even if linking them to care requires more effort.

Introduction

The last decade has seen a global transformation in HIV care, with the near universal availability of affordable and effective combination antiretroviral therapy (ART), that durably suppresses viral replication, prevents and even reverses progression to AIDS, and prevents

onward transmission of the virus. The discovery that combination ART blocks transmission of the virus led to the concept that annual testing followed by initiation of ART for infected individuals may be an effective form of HIV prevention [1].

The HPTN 071 (PopART) cluster-randomised trial evaluated whether a combination prevention package including universal testing and treatment would reduce HIV incidence in 21 communities in Zambia and South Africa [2]. In the home-based package, community workers conducted HIV testing and supported linkage to HIV care and ART adherence. The trial reported a 20% reduction of HIV incidence in study arms with the combination prevention package, but residual transmission remained.

The HPTN 071-02 (PopART) Phylogenetics Study was set up as an ancillary study to HPTN 071 (PopART) in the Zambian communities. The aims of the study were to use phylogenetics to find the origins and characterise the sources of ongoing transmission, assess the effectiveness of the intervention had it been rolled out nation-wide, and to identify the most promising policies for prevention in the future.

HIV phylogenies reflect the demographic and spatial history of transmission in a specific setting [3,4]. In population-level analyses, HIV phylogenetic inferences have been used to describe the origin and global spread of HIV [5], characterize transmission dynamics at the population level in concentrated epidemics [6], and, building on studies that examined individual transmission events [7], to analyse transmission linkages in HIV prevention trials [1,8,9]. More recently, phylogenetic analysis has been applied to understand epidemics in sub-Saharan African settings, looking at patterns of clustering [10] and spatial spread [11–13].

To understand the risk factors for onward HIV transmission, it is necessary to conduct large phylogenetic studies with high sampling density, such that probable source-recipient pairs can be studied. Two prior studies used this approach to identify the most likely transmitters and risk-factors for transmission in two well-characterized, concentrated European HIV networks consisting mostly of men who have sex with men [14,15]. Here, we report the first study of this kind in a large generalised epidemic in sub-Saharan Africa. A similar study is being conducted for the Ya Tsie Trial in Botswana [16]. We characterised the sources of transmission by age and gender, and determined if they were infected less than one year before the transmission, whether the sources were from a different community than the recipient, and if their dominant viral strain was resistant to first-line ART. Future work will

address the question of whether phylogenetic analysis reveals significant differences between the arms of the trial.

Methods

Study design

The HPTN 071 (PopART) cluster-randomised trial was conducted in Zambia and South Africa between 2013 and 2018 in 21 communities with a total population of approximately 1 million and adult HIV prevalence of 21.5% [2]. HPTN 071-02 (PopART) Phylogenetics was conducted in nine of the twelve Zambian PopART communities and included participants from the PopART Population Cohort (PC) and additional participants enrolled at health care facilities (HCF) in the same communities. The PC enrolled one randomly selected person aged 15-44 per randomly selected household, irrespective of HIV status, with a total of 20,264 participants in the nine Zambian communities between 2013 and 2018. At the HCF, participants were enrolled if they were HIV+ individuals aged 18 or over who started treatment (or restarted treatment after an interruption) between 2016 and 2018. Details can be found in the trial protocol (see https://www.hptn.org/sites/default/files/inline-files/HPTN%20071-2%2C%20Version%202.0%20%2807-14-2017%29.pdf).

Ethical considerations

Participants gave written consent specifically developed for the phylogenetics ancillary study (see protocol). Pseudonymised data were stored in a GDPR-compliant manner [17]. Ethical approvals for the HPTN 071 (PopART) trial and the HPTN 071-02 (PopART) Phylogenetics Study was granted by ethics committees at the London School of Hygiene and Tropical Medicine, University of Zambia, and Stellenbosch University, South Africa. The design of the study was discussed with community advisory boards. Results of the study were shared and discussed with participants in several local workshops.

Sequencing and phylogenetic analysis

Genetic sequences for the viral RNA in each sample were obtained using veSEQ-HIV, a bait-capture based quantitative high-throughput whole-genome deep-sequencing method with a sensitivity of >5,000 RNA copies/ml for whole genomes [18].

The phylogenetic analysis aimed to find phylogenetically probable source-recipient transmission pairs identified by analysing all obtained sequences. These methods analysed within-host genetic diversity assessed by deep sequencing data, building on the

phyloscanner software [19]. Analyses were performed using custom methods which have previously been validated against external reference datasets.

Phyloscanner was used to identify likely transmission pairs as described in Ratmann et al [20] from sequences aligned using *shiver* [21] and clusters generated using HIV-TRACE [22]. Same-sex pairs were excluded from further analysis as they were mostly predicted to be cases of indirect transmission with a missing intermediary (see supplement). The distribution of mixed-sex and same-sex pairs was not consistent with a significant contribution of cryptic transmission amongst men who have sex with men (see supplement). In order to estimate which transmissions had taken place in the trial period, the time since infection (TSI) was estimated for each sample using a machine learning approach (random forest regression) on the mean genomic diversity (minor allele frequency) and mean genomic divergence (largest subgraph root-to-tip distance) across all windows in the *gag*, *pol* and *env* genes. Direction of transmission in the pairs was reconstructed using two methods: a *phyloscanner* topological classification and comparison of the estimates of TSI for each individual. We included pairs called by one method and indeterminate in the other, and pairs called by both methods with concordant results. Sources were classified as coming from outside the community if they were sampled in a different study community than the recipient. To correct for sampling bias in all analyses, we calculated sample weights for the recipient population using iterative proportional fitting, according to four demographic variables: sex, birth cohort, community, and marital status.

Detection and classification of drug resistance mutations

A bioinformatic pipeline, drmSEQ, was used to predict drug resistance to first-line adult ART based on detection of mutations in the Illumina reads generated by veSEQ-HIV using the Stanford HIV Drug Resistance Database scoring system (HIVdb version 8.9.1) [23] as follows: scores 0-14 - WT/susceptible, scores 15-29 - low-level resistance, scores 30 and above - high-level resistance. The method has been previously validated against an FDA-approved drug resistance assay [24]. First-line adult ART national guidance at the time of sampling was non-nucleoside inhibitor efavirenz in combination with nucleoside inhibitors abacavir, AZT, D4T, DDI, FTC, 3TC or Tenofovir.

Results

Trial participants and sequences

Participants were enrolled from the PopART PC and HCFs in the same communities. Full or near-full sequence coverage was obtained for 5,613 participants (Figure 1, see detailed table in the supplement). Next generation sequencing (NGS) files are available via the PANGEA consortium [25] (www.pangea-hiv.org).

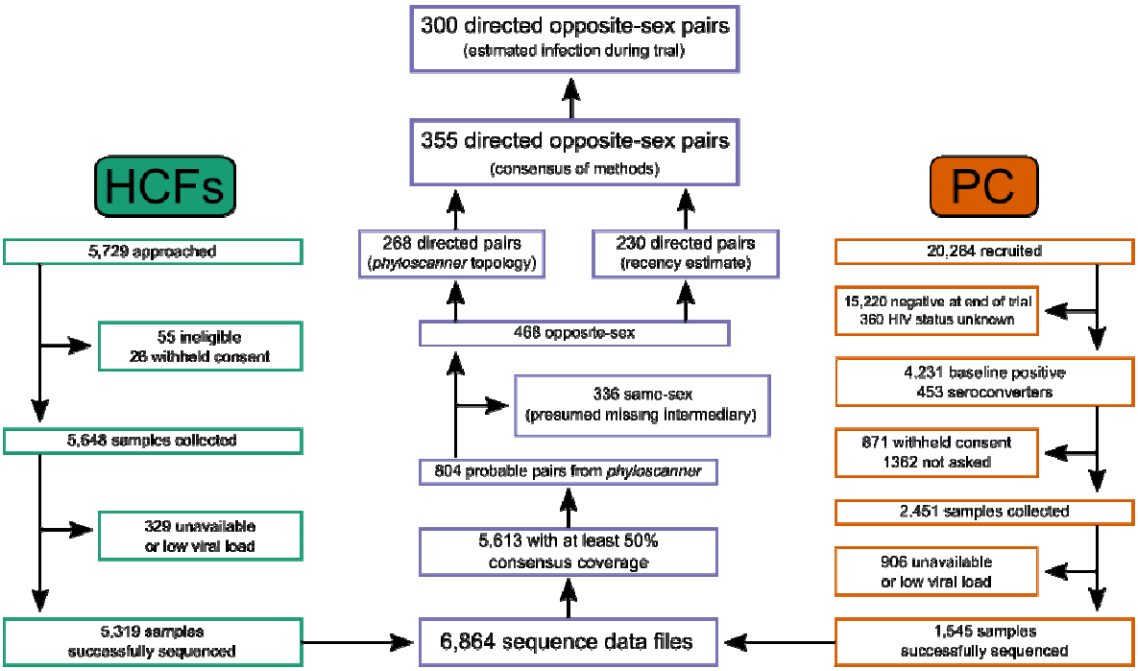


Figure 1: *Flowchart depicting determination of transmission pairs from recruited participants.*

Identification of transmission pairs

Phylogenetic analysis of all NGS reads identified 468 probable opposite-sex transmission pairs. A model of sampling of transmission chains estimates that each of these pairs has a probability of 90.4% of representing direct transmission (see supplement). Next to the 468 opposite-sex pairs, the analysis yielded 336 same-sex pairs, which were excluded from further analysis as they are likely to represent unsampled intermediaries (see supplement). The distribution of mixed-sex and same-sex pairs was not consistent with a significant contribution of cryptic transmission amongst men who have sex with men (see supplement). The likely direction of transmission in the pairs was assessed by two independent methods, using phylogenetic topology and estimates of recency of infection. Phylogenetic topology identified 268 directed pairs, the recency estimate 230 pairs. 143 pairs were called by both methods, 24 pairs were excluded due to conflicting result. The combined analysis yielded

355 probable transmission pairs with good confidence in the direction of transmission. For 300 of these, transmission was estimated to have occurred during the trial period (i.e., since January 2014) (Figure 1). An example of how a transmission pair is identified via topology is depicted in Figure 2. The number of transmission pairs identified was consistent with a power calculation conducted before the study to determine the number of transmission pairs required to characterise the sources of infection (see supplement).

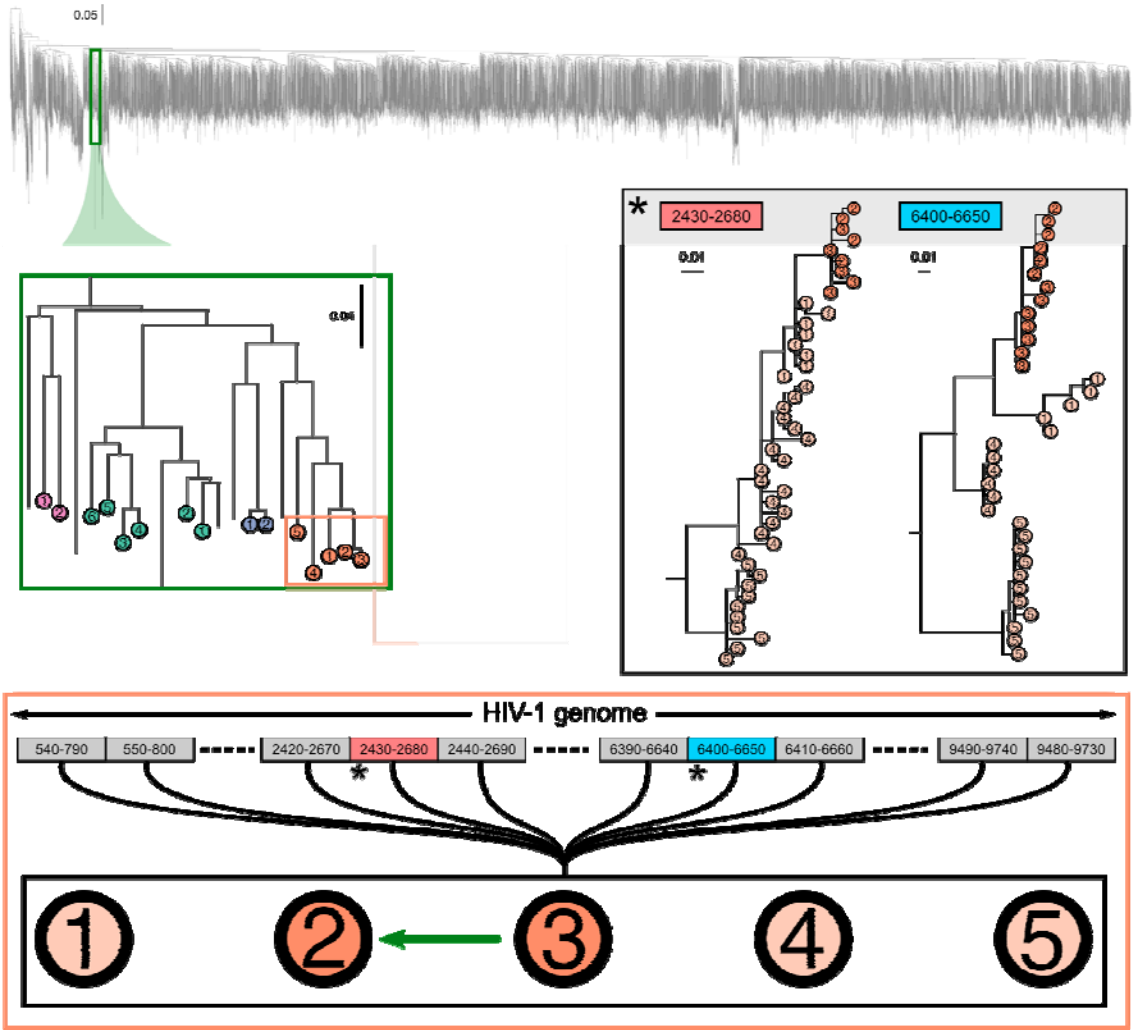


Figure 2: *Illustration of the selection of transmission pairs. The full phylogeny of the 5,613 consensus sequences with at least 50% coverage (top) has a subtree (green) enlarged. Within that subtree, coloured tip shapes indicate members of four separate HIV-TRACE clusters. A single cluster (orange) is further enlarged. A separate phylogeny of within-host tips was constructed for each of 898 genomic windows; two examples for windows 2430-2680 and 6400-6650 are presented (with reference sequences pruned). In these windows*

(asterisk, inset), and when all windows were summarised (bottom), orange host 3 was reconstructed as the source of the infection in orange host 2. The other three members of the orange cluster (1,4,5) were not identified as belonging to any directed transmission pair. Branch lengths in all phylogenies are in substitutions per site.

Characteristics of sources

We characterised the sources by gender, age, recency of infection, resistance to first-line ART, and whether they were registered in a different community than the recipients. We used iterative proportional fitting (raking) to weight the population of recipients so they were demographically representative of the population of all recipients who acquired HIV infection during the trial; proportions and medians given in this section are post-weighting.

The age distributions for both males and females showed the highest transmission from middle-aged individuals: 72.4% (95% CI: 65.6%-80.3%) of male sources were between 25 and 40 years of age at the time of transmission, representing 43.1% (36.6%-49.5%) of all transmissions; 72.8% (64.3%-81.3%) of female sources were between 20 and 35 years of age at the time of transmission, representing 29.5% (23.6%-35.5%) of all transmissions (Figure 3A). The median age at time of transmission was 32 years (range 20-64) for male sources and 25 (17-49) for female sources (Figure 3A). The distributions are wide and consistent with a generalised epidemic in which risk is ubiquitous and many individuals transmit to their regular partners.

After weighing, 59.4% (53.1%-65.8%) of transmissions were from male to female participants (Figure 3A). Since the adult HIV prevalence was 2.0 times higher in women than men, the per capita rate of transmission was thus 2.93 times higher for infected males than for infected females.

To determine at which stage of their infection sources were most likely to transmit, we used a machine learning approach on phylogenetic parameters to estimate how long the virus had been evolving in each host. This allowed us to estimate how long ago each individual had been infected, when the transmission event took place and for how long the source had been infected at transmission. In the transmission pairs, 72 recipients were infected by sources who themselves were estimated to have been infected within the 12 months before the transmission occurred (Figure 3D). This corresponds to a weighted proportion of 25.6% (19.9%-31.3%).

After omitting eleven individuals whose viruses had an unknown drug resistance profile, a large majority of sources (236/289) transmitted drug-sensitive viruses, while 20 participants transmitted HIV with low-level resistance to first-line ART and 33 HIV with high-level resistance to first-line ART (Figure 3E). The weighted proportions of individuals transmitting sensitive, low-level resistant and high-level resistant HIV were 82.5% (77.5%-87.5%), 6.7% (3.4%-10%) and 10.8% (6.7%-14.9%), respectively.

Finally, we determined the fraction of transmission pairs for which the source was living in a different community than the recipient. 263 of 300 transmissions occurred within the same community, a weighted proportion of 87% (83.6%-91.4%). For the interpretation of the HPTN 071 (PopART) trial, 13% cross-community transmission means that the effectiveness of the trial was likely underestimated in the original analysis, indicating that the PopART intervention package would have been more effective if it had been implemented country-wide.

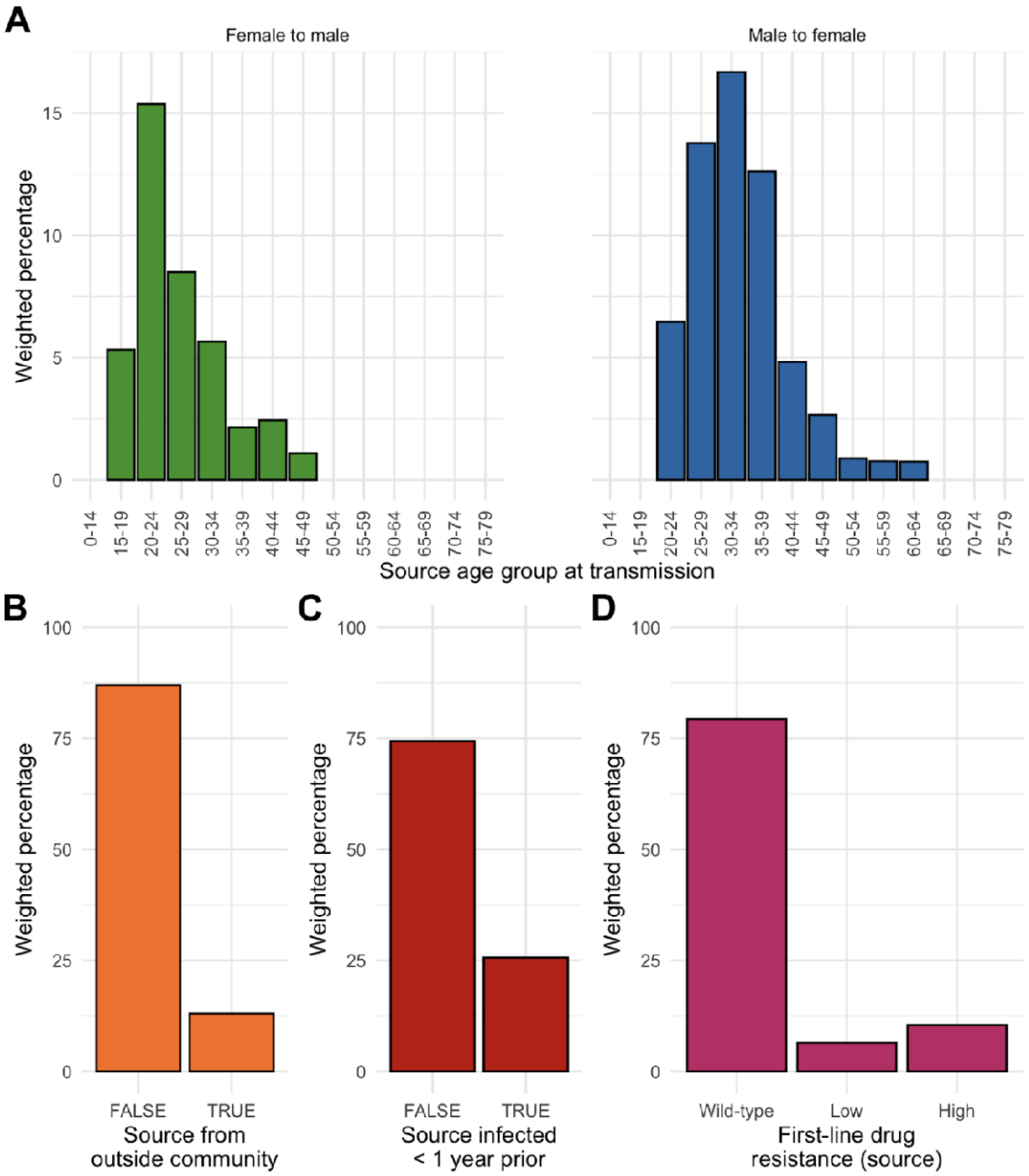
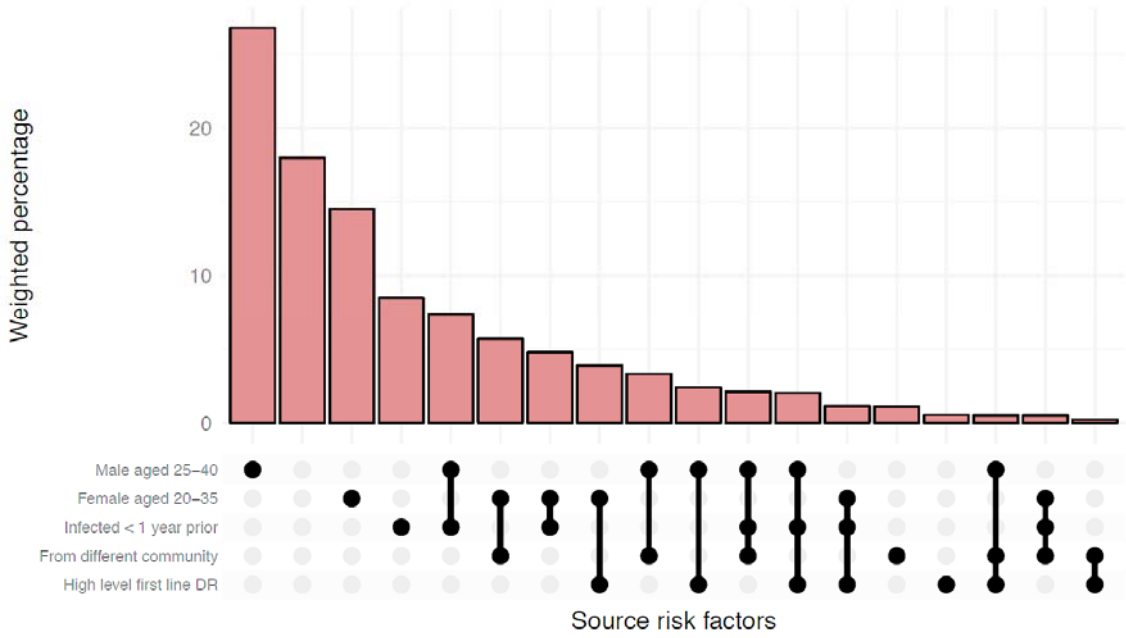


Figure 3: Histograms of characteristics of source individuals in the set of directed opposite-sex pairs determined to have been infected during the trial period. Counts have been weighted to make the recipient population representative of the population infected during the trial, using the PopART model for birth cohort, sex and community, and the 2018 Zambia demographic survey for marital status. A) Age profile of sources (at the estimated time of transmission). B) Infections caused by sources residing in a different community from the recipient. C) Recency of source infection (source estimated to be infected less than 1 year before recipient). D) First-line drug resistance profiles.

Combined characteristics of sources

To gain a better understanding of how the different characteristics of sources interact with each other, we ranked them according to the fraction of transmission in which they played a role (Figure 4A). The highest fraction of transmissions originated from males aged 25-40 who were infected more than a year prior to transmission and lived in the same community. They were the sources for twice as many transmissions as the female risk group aged 20-35 with the same characteristics and almost twice as many transmissions as all sources outside these two age groups. These groups were followed by the same groups who were infected less than a year before transmission, were registered in a different community, and whose dominant viral population was resistant to first-line ART. There was no group that disproportionately combined several of the risk factors, as real values were generally very close to values expected if risk factors present in the population of sources had been randomly assigned to this population (Figure 4B).

A



B

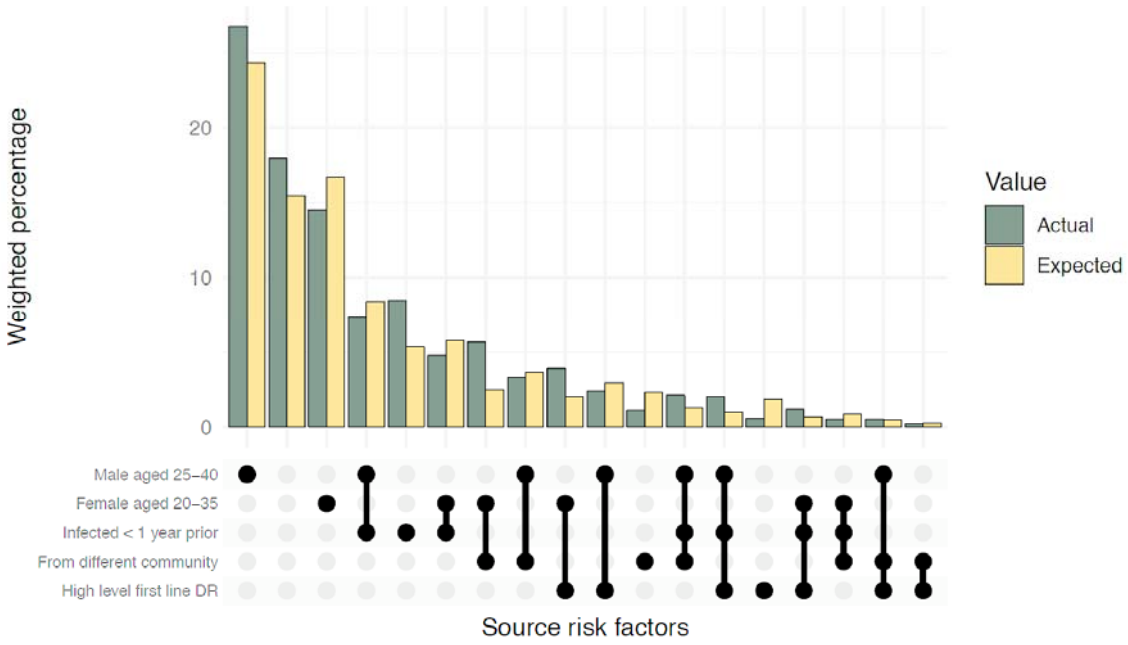


Figure 4: *A) Distribution of all combinations of five key risk factors amongst the set of directed opposite-sex pairs determined to have been infected during the trial period. Each bar represents a group of sources whose characteristics are defined by the black dots in the bottom rows. Bars are weighted as in previous plots. B) Actual distribution of risk factors compared to the distribution expected if risk factors had been allocated randomly to sources based on risk factor percentages identified in the source population.*

The analysis also shows which transmissions could have been averted by the PopART intervention package and which would have been missed. Testing in the trial was approximately yearly which means that transmissions caused by groups which have a black dot in the third row might have been missed. Transmissions from a different community would likely also have been missed if they originated in one of the control communities in which the prevention package was not offered. The analysis offers a convenient way to assess different prevention policies and the maximum impact on incidence they are likely to have.

Discussion

This study identified likely transmission pairs and the characteristics of sources of HIV transmission in Zambia between 2014 and 2019. The aim of the study is to provide policy makers with data on where transmission is still occurring in a generalised epidemic in sub-

Saharan Africa despite delivery of a universal testing and treatment intervention and advise where prevention efforts should be focussed.

We identified men aged 25-40 as the group most commonly transmitting HIV. This group was responsible for 43.2% of transmissions. Women aged 20-35 were responsible for 30.3% of transmissions. Age distributions of sources were wide, suggesting that risk of infection was common across a large fraction of society. We did not find evidence for a renewal cycle involving chronically infected older women as suggested previously [26]. As prevalence among women was 2.0 times as high as prevalence among men in the HPTN 071 (PopART) study, the number of new infections from young men per infected young man was 2.93 times the same number for young women. Men in this age group were less likely than other demographic groups to be linked to care and on ART [27]. These results highlight the need for more intensive efforts in UTT programmes to achieve high coverage of men aged 25-40 and suggest that pre-exposure prophylaxis services should also be considered, particularly for young women [28,29].

A quarter of sources in this study were estimated to have been infected for less than a year before passing on the infection. Previous estimates have ranged between 5-10% and 60-80%, reviewed in [30,31]. Our results suggest that both recently infected individuals and individuals with chronically unsuppressed HIV require attention: more frequent testing to avert transmission from recent infections and efforts to retain people in care to avert transmission from chronic infection.

In this study, 10.4% and 6.4% of sources' viral population were dominated by variants with high-level and low-level resistance against first-line ART, respectively. This mirrors the high percentage of first-line ART resistance in Sub-Saharan Africa reported by other studies in the past few years [32–34]. The rollout of dolutegravir is urgently needed and will hopefully reverse this trend. Resistance against first-line ART including dolutegravir should be monitored by local surveillance teams as availability of HIV drug resistance testing and NGS sequencing in sub-Saharan Africa increases.

A large majority of transmissions occurred between partners living in the same community. However, 13% of transmissions occurred between partners living in different communities. As the primary analysis of the HPTN 071 (PopART) trial [2] took no account of between-community transmission, our findings indicate that HIV incidence would have been further reduced if the PopART prevention package had been implemented country-wide. The complex statistical corrections required to assess the impact of the intervention taking into

account between-community transmission are outside the scope of this study, but further analysis of the phylogenetics data is likely to lead to a higher estimate of the effectiveness of the intervention.

This study has limitations. While sampling fractions have been high and reached between 4.9% and 9.5% of people living with HIV in the individual study communities, the sample was still not fully representative of the HIV-positive population as a whole. In particular, recruitment of one person per household in the PC limited our ability to identify within-household transmission pairs. This is partially mitigated by the collection of samples in the HCFs which included additional members of the same households. The sequencing approach yielded sequences for participants with more than 5000 viral copies/mL. This corresponds to 92.8% of participants with viral loads above 1500 copies/mL who might have been involved in a transmission event [35].

HPTN 071-02 (PopART) Phylogenetics is the largest HIV phylogenetic study conducted to date, the first large transmission study to be based on an a priori power calculation, and the the most comprehensive study of characteristics of sources of HIV-1 infection in sub-Saharan Africa. The study highlights that residual transmission occurred across a wide spectrum of common sexual interactions and was not limited to small risk groups. Men aged 25-40 however were, as a group, responsible for a large share of transmissions and should be prioritised in prevention efforts, even if linking them to care requires more effort.

Study Sponsorship and Funding statement

HPTN 071 is sponsored by the National Institute of Allergy and Infectious Diseases (NIAID) under Cooperative Agreements UM1-AI068619, UM1-AI068617, and UM1-AI068613, with funding from the US President's Emergency Plan for AIDS Relief (PEPFAR). Additional funding is provided by the International Initiative for Impact Evaluation (3ie) with support from the Bill & Melinda Gates Foundation, as well as by NIAID, the National Institute on Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH), all part of NIH. HPTN 071-02 Phylogenetics is sponsored by NIAID, NIMH, and the Bill & Melinda Gates Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIAID, NIMH, NIDA, PEPFAR, 3ie, or the Bill & Melinda Gates Foundation.

Author contributions

Matthew Hall - conducted analysis, co-wrote paper
Tanya Golubchik - conducted analysis, developed method, edited paper
David Bonsall - conducted analysis, developed method, managed laboratory, edited paper
Lucie Abeler-Dörner - co-wrote paper, management
Mohammed Limbada - clinical site management, training, edited paper
Barry Kosloff - laboratory site management, training, edited paper
Ab Schaap - data management
Mariateresa de Cesare - sequenced viruses
George Mackintyre-Cockett - sequenced viruses
William Probert - conducted analysis
Oliver Ratmann - developed protocol
Ana Bulas Cruz - project management
Estelle Piwowar-Manning - site laboratory auditing
David Burns - developed protocol, edited paper
Myron Cohen - developed protocol, edited paper
Deborah Donnell - population cohort data management
Sue Eshleman - laboratory centre management, edited paper
Musonda Simwinga - community engagement, dissemination of findings
Helen Ayles - site principal investigator, managed study team, dissemination of findings, edited paper
Sarah Fidler - co principal investigator of HPTN 071 and 071-2, clinical oversight, protocol development, edited paper
Richard Hayes principal investigator of HPTN 071, edited paper
Christophe Fraser principal investigator, developed protocol, conducted analysis, managed study, co-wrote paper

References

1. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365(6):493–505.

2. Hayes RJ, Donnell D, Floyd S, et al. Effect of Universal Testing and Treatment on HIV Incidence - HPTN 071 (PopART). N Engl J Med 2019;381(3):207–18.

3. Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC. Human immunodeficiency virus. Phylogeny and the origin of HIV-1. Nature 2001;410(6832):1047–8.

4. Gray RR, Tatem AJ, Lamers S, et al. Spatial phylodynamics of HIV-1 epidemic emergence in east Africa. AIDS 2009;23(14):F9–17.

5. Faria NR, Rambaut A, Suchard MA, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science 2014;346(6205):56–61.

6. Kouyos RD, von Wyl V, Yerly S, et al. Molecular epidemiology reveals long-term changes in HIV type 1 subtype B transmission in Switzerland. J Infect Dis 2010;201(10):1488–97.

7. Leitner T, Escanilla D, Franzén C, Uhlén M, Albert J. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. Proc Natl Acad Sci U S A 1996;93(20):10864–9.

8. Eshleman SH, Hudelson SE, Redd AD, et al. Analysis of genetic linkage of HIV from couples enrolled in the HIV Prevention Trials Network 052 trial. J Infect Dis 2011;204(12):1918–26.

9. Campbell MS, Mullins JI, Hughes JP, et al. Viral linkage in HIV-1 seroconverters and their partners in an HIV-1 prevention clinical trial. PLoS One 2011;6(3):e16986.

10. Novitsky V, Bussmann H, Logan A, et al. Phylogenetic relatedness of circulating HIV-1C variants in Mochudi, Botswana. PLoS One 2013;8(12):e80589.

11. Novitsky V, Zahralban-Steele M, Moyo S, et al. Mapping of HIV-1C Transmission Networks Reveals Extensive Spread of Viral Lineages Across Villages in Botswana Treatment-as-Prevention Trial. J Infect Dis 2020;222(10):1670–80.

12. Ratmann O, Kagaayi J, Hall M, et al. Quantifying HIV transmission flow between high-prevalence hotspots and surrounding communities: a population-based study in Rakai, Uganda. Lancet HIV 2020;7(3):e173–83.

13. Bbosa N, Ssemwanga D, Ssekagiri A, et al. Phylogenetic and Demographic Characterization of Directed HIV-1 Transmission Using Deep Sequences from High-Risk and General Population Cohorts/Groups in Uganda. Viruses [Internet] 2020;12(3). Available from: http://dx.doi.org/10.3390/v12030331

14. Fisher M, Pao D, Brown AE, et al. Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. AIDS 2010;24(11):1739–47.

15. Ratmann O, van Sighem A, Bezemer D, et al. Sources of HIV infection among men having sex with men and implications for prevention. Sci Transl Med 2016;8(320):320ra2.

16. Lockman S, Holme MP, Makhema J, et al. Implementation of Universal HIV Testing and Treatment to Reduce HIV Incidence in Botswana: the Ya Tsie Study. Curr HIV/AIDS Rep 2020;17(5):478–86.

17. Mondschein CF, Monda C. The EU's General Data Protection Regulation (GDPR) in a Research Context. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science. Cham (CH): Springer; 2018.

18. Bonsall D, Golubchik T, de Cesare M, et al. A Comprehensive Genomics Solution for HIV Surveillance and Clinical Monitoring in Low-Income Settings. J Clin Microbiol [Internet] 2020;58(10). Available from: http://dx.doi.org/10.1128/JCM.00382-20

19. Wymant C, Hall M, Ratmann O, et al. PHYLOSCANNER: Inferring Transmission from Within- and Between-Host Pathogen Genetic Diversity. Mol Biol Evol 2018;35(3):719–33.

20. Ratmann O, Grabowski MK, Hall M, et al. Inferring HIV-1 transmission networks and sources of epidemic spread in Africa with deep-sequence phylogenetic analysis. Nat Commun 2019;10(1):1411.

21. Wymant C, Blanquart F, Golubchik T, et al. Easy and accurate reconstruction of whole HIV genomes from short-read sequence data with shiver. Virus Evol 2018;4(1):vey007.

22. Kosakovsky Pond SL, Weaver S, Leigh Brown AJ, Wertheim JO. 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(7):1812–9.

23. Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis 2006;42(11):1608–18.

24. Fogel JM, Bonsall D, Cummings V, et al. Performance of a high-throughput next-generation sequencing method for analysis of HIV drug resistance and viral load. J Antimicrob Chemother 2020;75(12):3510–6.

25. Abeler-Dörner L, Grabowski MK, Rambaut A, Pillay D, Fraser C, PANGEA consortium. PANGEA-HIV 2: Phylogenetics And Networks for Generalised Epidemics in Africa. Curr Opin HIV AIDS 2019;14(3):173–80.

26. de Oliveira T, Kharsany ABM, Gräf T, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. Lancet HIV 2017;4(1):e41–50.

27. Floyd S, Shanaube K, Yang B, et al. HIV testing and treatment coverage achieved after 4 years across 14 urban and peri-urban communities in Zambia and South Africa: An analysis of findings from the HPTN 071 (PopART) trial. PLoS Med 2020;17(4):e1003067.

28. Irungu EM, Baeten JM. PrEP rollout in Africa: status and opportunity. Nat Med 2020;26(5):655–64.

29. Grant H, Gomez GB, Kripke K, et al. Time to Scale Up Preexposure Prophylaxis Beyond the Highest-Risk Populations? Modeling Insights From High-Risk Women in Sub-Saharan Africa. Sex Transm Dis 2020;47(11):767–77.

30. Cohen MS, Dye C, Fraser C, Miller WC, Powers KA, Williams BG. HIV treatment as prevention: debate and commentary--will early infection compromise treatment-as-prevention strategies? PLoS Med 2012;9(7):e1001232.

31. Rutstein SE, Ananworanich J, Fidler S, et al. Clinical and public health implications of acute and early HIV detection and treatment: a scoping review. J Int AIDS Soc 2017;20(1):21579.

32. Rhee S-Y, Blanco JL, Jordan MR, et al. Geographic and temporal trends in the molecular epidemiology and genetic mechanisms of transmitted HIV-1 drug resistance: an individual-patient- and sequence-level meta-analysis. PLoS Med 2015;12(4):e1001810.

33. Gupta RK, Gregson J, Parkin N, et al. HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries: a systematic review and meta-regression analysis. Lancet Infect Dis 2018;18(3):346–55.

34. Chimukangara B, Lessells RJ, Rhee S-Y, et al. Trends in Pretreatment HIV-1 Drug Resistance in Antiretroviral Therapy-naive Adults in South Africa, 2000-2016: A Pooled Sequence Analysis. EClinicalMedicine 2019;9:26–34.

35. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med 2000;342(13):921–9.