

Published in final edited form as:

*Sci Transl Med.* 2016 January 6; 8(320): 320ra2. doi:10.1126/scitranslmed.aad1863.

## Sources of HIV infection among men having sex with men and implications for prevention \*

O. Ratmann<sup>1</sup>, A. van Sighem<sup>2</sup>, D. Bezemer<sup>2</sup>, A. Gavryushkina<sup>3</sup>, S. Jurriaans<sup>4</sup>, A. Wensing<sup>5</sup>, F. de Wolf<sup>1</sup>, P. Reiss<sup>2,6</sup>, and C. Fraser<sup>1</sup>

<sup>1</sup>Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom <sup>2</sup>Stichting HIV Monitoring, Amsterdam, the Netherlands

<sup>3</sup>Department of Computer Science, University of Auckland, New Zealand <sup>4</sup>Department of Medical Microbiology, Academic Medical Center, Amsterdam, The Netherlands <sup>5</sup>Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, the Netherlands <sup>6</sup>Department of Global Health, Academic Medical Center, Amsterdam, the Netherlands

### Abstract

New HIV diagnoses among men having sex with men (MSM) have not decreased appreciably in most countries, even though care and prevention services have been scaled up substantially in the past twenty years. To maximize the impact of prevention strategies, it is crucial to quantify the sources of transmission at the population level. We used viral sequence and clinical patient data from one of Europe's nation-wide cohort studies to estimate probable sources of transmission for 617 recently infected MSM. 71% of transmissions were from undiagnosed men, 6% from men who had initiated antiretroviral therapy (ART), 1% from men with no contact to care for at least 18 months, and 43% from those in their first year of infection. The lack of substantial reductions in incidence amongst Dutch MSM is not a result of ineffective ART provision or inadequate retention in care. In counterfactual modeling scenarios, 19% of these past cases could have been averted with current annual testing coverage and immediate ART to those testing positive. 66% of these cases could have been averted with available antiretrovirals (immediate ART provided to all MSM testing positive, and pre-exposure antiretroviral prophylaxis taken by half of all who test negative for HIV), but only if half of all men at risk of transmission had tested annually. With increasing sequence coverage, molecular epidemiological analyses can be a key tool to direct HIV prevention

---

**Corresponding author:** Oliver Ratmann [oliver.ratmann@imperial.ac.uk](mailto:oliver.ratmann@imperial.ac.uk).

\*This manuscript has been accepted for publication in Science Translational Medicine. This version has not undergone final editing. Please refer to the complete version of record at [www.sciencetranslationalmedicine.org/](http://www.sciencetranslationalmedicine.org/). The manuscript may not be reproduced or used in any manner that does not fall within the fair use provisions of the Copyright Act without the prior, written permission of AAAS.

#### Author contributions

OR, FW, PR, CF conceived the study. OR, CF developed the methods, did the analysis, and reviewed all statistical aspects of the analysis. AS, DB, SJ, AW provided data used to conduct the analysis. AG assisted in estimating the viral phylogeny. AS, DB, FW, PR advised on analysis and interpretation. OR, CF wrote the first draft. All authors reviewed and approved the final version.

**Competing interests:** None declared.

**Data and materials availability:** Data are available from the HIV Monitoring Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data. Contact email: [secretariaat.shm@amc.uva.nl](mailto:secretariaat.shm@amc.uva.nl).

strategies to the predominant sources of infection, and help send HIV epidemics amongst MSM into a decisive decline.

---

## Introduction

Combination antiretroviral therapy (ART) transformed HIV from a deadly to a life-long disease, and is also one of the most effective strategies for preventing onward infections (1, 2). However, among men having sex with men (MSM), the substantial scale-up of ART in the past twenty years has not resulted in appreciable reductions of new HIV infections and diagnoses (table 1) (3). Building on successful behavioural and biomedical HIV prevention strategies (4), further interventions exist that could be used to reduce the number of HIV infections amongst MSM. The 2016 WHO guidelines now recommend initiation regardless of CD4 cell count after diagnosis (immediate ART), as well as provision of antiretrovirals as pre-exposure prophylaxis (PrEP) to those at substantial risk of infection (5). Future prevention programmes could focus on one or both recommended interventions, as well as on increased routine HIV testing and diagnosis (6); RNA testing to detect MSM in early acute infection when they are thought to be the most infectious (7); and improved adherence and linkage support to assist patients with attaining and sustaining undetectable viral loads whilst on ART (8). The potential impact of any of these interventions, and specifically those recommended by the WHO, relies crucially on how many HIV transmissions originate from different stages in the entire HIV infection and care continuum, ranging from undiagnosed acute infection through treated infection and loss to follow-up. This has been challenging to measure directly through classical epidemiological approaches.

In this study, we use the viral phylogenetic relationship between partial HIV-1 subtype B polymerase sequences to reconstruct past, probable transmission events in the Netherlands (figure 1). These sequences were routinely collected for drug resistance testing of HIV-infected patients that are in care (9). Amongst sampled MSM, 94% were of subtype B. Then, we use clinical records to determine the staging of probable transmission events within the infection and care continuum (figure 2A and table 2). This enabled us to estimate the population-level proportion of transmissions amongst the reconstructed transmission events that are attributable to fourteen stages of the infection and care continuum in figure 2A. Transmissions could be attributed to stages before diagnosis because HIV sequences, always collected after diagnosis, diverge fast enough to indicate past transmission events (10). Similarly, transmissions could also be attributed to men with no contact to care for at least 18 months. Finally, using these estimates, we quantified the potential impact of available, but currently not implemented prevention programmes in the Dutch MSM population, had they been used in the last three years. In particular, we evaluate if the revised 2016 WHO guidelines on immediate ART and PrEP could have substantially altered the course of the Dutch HIV epidemic amongst MSM.

Understanding which interventions should be prioritized for the Dutch MSM epidemic is an important case study. First, the number of new MSM infections in the Netherlands has not decreased appreciably (9) despite comprehensive linkage and retention in care, substantial ART scale up free of charge, and frequent follow up to maintain viral control of the vast

majority of those on ART (table 1). Second, similar epidemic trends are reported from other countries with an overall equally comprehensive cascade of care (table 1), casting more general doubts on the population-level impact of current prevention strategies targeting MSM epidemics (11). Third, nearly all HIV-infected MSM in care are enrolled in the clinical, national opt-out ATHENA cohort since early 1996 (12). HIV care is monitored comprehensively at high frequency (clinic visits, treatment histories, co-morbidities recorded; ~3 viral load/CD4 measurements per year per individual) (12), which allowed us to characterize phylogenetically reconstructed transmission events in detail.

## Results

### Potential transmissions to MSM in confirmed recent infection at time of diagnosis

By 2013, 11,863 HIV-infected MSM were registered and still in care in the Netherlands. To estimate their sources of transmission and then the impact of prevention programmes, we focussed on transmissions to MSM that were recently infected at time of diagnosis (stage A in figure 1). Between July 1996 and December 2010, 1,794 MSM had been infected at most 12 months prior to diagnosis. Types of evidence were a previous negative HIV test (76%), laboratory diagnosis (7%), or clinical diagnosis of acute infection (17%). For 1,045 (58%) of these, a sequence was available. To these recipient MSM, we considered as potential transmitters all HIV-infected men whose course of infection overlapped with the infection window of the recipient (stage A in figure 1). With this approach, we could resolve the timing and direction of potential transmission events (13). Out of all 12,207 potential transmitters, 5,593 (46%) had a viral sequence and formed ~ 4.4 million potential transmission pairs with sequences available for both individuals (stage B in figure 1).

### Phylogenetically probable transmission events

Genetic sequences of the virus alone cannot prove epidemiological linkage (14). However, most of the potential transmission pairs could be ruled out as implausible, based on the phylogenetic relationship of the viral sequences. The viral phylogeny among the Dutch sequences and their closest matches in the Los Alamos HIV sequence database (<http://www.hiv.lanl.gov/>) was reconstructed with maximum-likelihood methods, and reliable subtrees were identified (see Material and Methods). Potential transmitters whose sequences did not occur in the same reliable subtree as those of the recipient MSM were excluded (stage C in figure 1) (14), as were potential transmitters whose sequences were incompatible with a direct HIV transmission event (stage D in figure 1) (15). Direct transmission could be excluded in 99.96% of all potential transmission pairs. We identified 903 phylogenetically probable transmitters to 617 recipient MSM in 2,343 pairs. Our analyses are based on this open observational cohort of past, phylogenetically reconstructed transmission events.

To guide and interpret this exclusion analysis, we evaluated patterns of viral divergence between sequences isolated from epidemiologically confirmed transmission pairs (16), and pairings of Dutch MSM that could not have infected each other (see Material and Methods). Based on these pairs, the above exclusion criteria were highly specific (true transmitters to recipients are not excluded, >90%), whilst sensitivity was low (incorrect transmission pairs could not always be excluded, ~60%). This indicates that the actual transmitter is almost

certainly among the phylogenetically reconstructed, probable transmitters, provided he was sequenced. From the known sequence coverage alone, we expected that approximately half of all 1,045 recipient MSM with a sequence had their actual transmitter sampled— suggesting further that the actual transmitter is among the phylogenetically reconstructed, probable transmitters for the large majority of the reconstructed 617 transmission events.

Clinical and demographic characteristics of the selected 617 recipient MSM were typical of all 1,794 MSM that were in confirmed recent infection at time of diagnosis (table 3). This indicates that the probable transmitters in the cohort are also typical of the transmitters to recently infected MSM.

### **Characterization of individual transmission events by stage in the HIV infection and care continuum**

Using clinical records, we then enumerated all stages in the HIV infection and care continuum during which the 617 transmission events could have occurred. Probable transmitters progressed in stage over time, and overlapped with infection windows in 13,169 time-resolved, six week long transmission intervals (figure 2B). Censoring and sequence sampling biases were identified for each stage by comparing men with and without a sequence, and were adjusted in line with previous work (17). Reflecting targeted sequence collection, intervals were not missing at random (figures 2C and S9). Each interval was associated with a phylogenetic transmission probability, based on the genetic distance between sequences from the transmitter and recipient and the time elapsed since the putative transmission interval and the sampling dates of both individuals (see Materials and Methods and figure S10). For each recipient, the probability that transmission occurred from one of the fourteen stages then depends on the number of his probable transmitters in that stage, and the transmission probabilities associated with each of the corresponding transmission intervals (see Materials and Methods).

### **Sources of HIV transmission**

The population-level proportions of HIV transmissions attributable to the fourteen infection/care stages were obtained by summing individual-level transmission probabilities by stage across all recipients, and are shown in table 4. Figure 3 compares the proportion of transmissions from each stage to the population-level proportion of infected men in these stages. Between July 1996 and December 2010, an estimated 71% [66%-73%] of all 617 transmission events originated from undiagnosed men, 22% [21%-26%] from diagnosed but not yet treated men, 6% [5%-8%] from men who initiated ART and 1% [0.7-1.6%] from men with no contact to care for at least 18 months. An estimated 43% [37%-46%] of the 617 recipient MSM were infected by men undergoing their first year of infection.

### **Impact of prevention strategies**

Figure 4 describes the counterfactual prevention scenarios for which we calculated the proportion of transmissions in the cohort that could have been averted between mid 2008 to December 2010, had we intervened to re-distribute the identified, probable transmitters to less infectious infection/care stages. Young MSM are at particularly high risk of infection (18, 19). We therefore considered— along the revised 2016 WHO guidelines (5)— roll-out

of immediate ART to all infected MSM and PrEP to half of all MSM aged 30 or less that test negative: at most 30% [22%-39%] of infections could have been averted without increased annual testing. Immediate ART alone could have averted 19% [13%-26%] of these cases at current testing levels. In practice, low adherence is associated with decreasing effectiveness of PrEP (20). We assumed an 86% efficacy of PrEP as reported in the recent Ipergay and PROUD trials (21, 22). Figure S12 reports the impact of lower efficacy values. Figure S13 reports the impact of lower or higher PrEP coverage. Next, we considered increased annual testing. Only 17% of identified probable transmitters had a last negative test in the year before diagnosis, compared to 27% of diagnosed MSM between mid 2008 to December 2010 and 38% of uninfected MSM in 2013 (table 1). If half of all transmitters had tested annually, immediate ART and PrEP to half of all MSM aged 30 or less that test negative could have averted 45% [34%-56%] of infections. Additional roll-out of PrEP to half of all men testing negative would have substantially boosted the combination intervention: 66% [50%-78%] of infections could have been averted.

## Discussion

HIV epidemics amongst MSM have—unlike other settings (23)—not declined appreciably with substantial improvements to care and ART scale-up (table 1). We characterized 617 past transmission events amongst MSM in the Netherlands based on phylogenetic and clinical data, estimated their sources throughout the infection and care continuum, and quantified the impact that biomedical prevention programmes could have had in averting the reconstructed transmission events. Analysing this transmission cohort, we aim to inform the design of future prevention interventions beyond high levels of ART coverage and the numerous successful behavioural interventions that are already in place (9).

A potential limitation of this study is that transmitters to MSM in recent infection at diagnosis may differ from typical transmitters. On average, fewer men diagnosed late with a CD4 count below 350 cells/ml occurred in phylogenetic transmission clusters with a recipient MSM, compared to those without (figure S23). This may imply that overall, the proportion of transmissions from undiagnosed men in chronic infection is higher, and consequently that the impact that immediate ART could have had is lower than our estimates. Conversely, the impact of increased annual testing and PrEP could be larger than reported, if men diagnosed late are not more difficult to reach than the average transmitter in our cohort. Further, this study focuses on the sources and prevention of in-country transmissions: 97% of the recipient MSM reported that infection was likely acquired in the Netherlands compared to 86% of diagnosed MSM. The contribution of cross-border transmissions may increase as the response is strengthened (24), an effect which we did not consider. Phylogenetic uncertainty and the phylogenetic exclusion criteria had little impact on our findings (figures S14-S22). A further potential caveat to the robustness of our findings is that only half of all potential transmitters had a viral sequence sampled. Although population-level sampling biases were adjusted, we must acknowledge that the actual transmitter may not have been sampled for all recipients. Improving sequence sampling coverage at time of diagnosis is needed to facilitate phylogenetic prevention analyses (25).

The identified sources of transmission imply, first, that viral suppression induced by ART is highly effective in preventing transmissions in this population (figure 3). The relative risk of HIV transmission from men after ART initiation varies by stage but is always estimated well below one when compared to diagnosed, untreated men with a CD4 count above 500 cells/ml, and is in particular 0.04 [0.02-0.1] for men with viral suppression (figure S11).

Second, very few transmissions are attributable to temporary or permanent loss to follow up, which must be considered in the context of high linkage and retention to care in the Netherlands: few diagnosed MSM had subsequently no contact to care for at least 18 months (8.2%) and most reentered care within five years (69%) (9). In contrast, several studies indicate that more than half of all transmissions amongst MSM in the United States originate from men that were not retained in care (26–28). The estimated impact of particular prevention strategies in figure 4 is limited to settings with a similar epidemic profile and care cascade as the Netherlands (table 1).

Third, not more than an estimated 20% of infections in the cohort could have been averted between mid 2008 and December 2010 with immediate ART after diagnosis. Given the remarkable expansion of ART coverage in the Netherlands in the past (9), the prevention potential of immediate ART is now limited. Nonetheless, starting ART at a cell count above 500 cells/ml leads to improved clinical outcomes and remains a priority (29).

Fourth, and similar to other locations (25, 30), almost half of all infections in our transmission cohort originated from men in their first year of infection. Frequent early transmission limits the overall impact of annual testing plus immediate ART to those testing positive (figure 4), and implies that prevention services to uninfected MSM must be strengthened. The substantial, estimated impact that PrEP would have had in averting transmissions in our cohort (figure 4) supports making PrEP available to MSM testing negative as in the United States (31). Recent PrEP demonstration projects (32, 33) indicate that existing barriers such as low awareness (34) and a lack of experience amongst providers (35) can be addressed. Concerns regarding the toxicity of PrEP, increasing sexual risk behaviour and emerging drug resistance have to date not been substantiated since PrEP was made available in the United States (36). In the context of PrEP-experienced prevention services, high discontinuation rates after PrEP initiation appear to be the greatest challenge to maintain protection from infection (32).

Fifth, without substantial increases to current annual testing coverage, ART and PrEP offered along the revised 2016 WHO guidelines could not have prevented more than a quarter of all infections in our transmission cohort. Since phylogenetically probable transmitters tend to test much less frequently than the average diagnosed MSM, substantial barriers likely exist in reaching men at high risk of onward transmission, and further work is needed to characterize these (37). Strategies such as self-testing (38), community-based testing (39), and more provider-initiated routine testing in general practices and at medical admissions raised annual testing coverage in pilot projects (12), and need to be expanded alongside biomedical interventions.



Sixth, this study indicates that substantial reductions in HIV incidence amongst MSM could have been realized with a combination approach that includes—critically—increased annual testing, with uptake of PrEP by young MSM testing negative and provision of immediate ART to those testing positive. This finding is primarily based on the impact of increased annual testing and the higher efficacy of PrEP reported in two recent randomized controlled trials (21, 22), and updates previous studies that estimate more limited benefits (4, 40, 41). Beyond age at testing, other characteristics not available to this study may also indicate high infection risk (42), and thereby identify groups of MSM to which PrEP should be made available as a priority. Provision of PrEP to all men testing negative is not affordable at current drug prices in high-income countries (40). The magnitude of the predicted impact of test-and-PrEP-and-treat for all (figure 4) could set an aspirational target for the fight against HIV amongst MSM.

The lack of substantial reductions in incidence amongst Dutch MSM is not a result of ineffective ART provision or inadequate retention in care. New HIV infections amongst MSM are challenging to prevent due to frequent early transmission and continued low testing uptake of men at risk of transmission. Counterfactual prevention scenarios on phylogenetically reconstructed, past transmission events to MSM in recent infection at diagnosis predict that increased annual testing and uptake of PrEP by men at high risk of infection have a key role to send the HIV epidemic amongst MSM into a decisive decline.

## Materials and Methods

### Study design

We conducted a retrospective viral phylogenetic transmission and prevention study that focuses on transmissions to MSM in confirmed recent HIV infection at time of diagnosis in the Netherlands (figure 1). The pre-specified objectives were to, first, reconstruct past, phylogenetically probable transmission events to these recipient MSM; second, to estimate the proportion of transmissions originating throughout the infection and care continuum based on the reconstructed transmission events; and, third, to estimate the proportion of infections that could have been averted through reallocating past, probable transmitters to less infectious stages in counterfactual modeling scenarios.

The ATHENA national observational HIV cohort includes anonymized data of all HIV-infected patients followed longitudinally in the 27 HIV treatment centres in the Netherlands since 1996, except 1.5% who opt-out (9). ATHENA patients are informed of data collection by their treating physician and can refuse further collection of clinical data according to an opt-out procedure. Patients who were diagnosed between 1981 and 1995 were included in the cohort when they were still alive in 1996 (9). Demographic, clinical, and viral sequence data were collected at entry and follow-up visits as described previously (9). By March 2013, viral sequence data had been systematically entered until December 2010. Therefore, recipients were enrolled between early 1996 and December 2010. Potential transmitters were enrolled until database closure in March 2013. Table S1 characterizes the demographic, clinical, and viral sequence data that were used in this study. The resolution of the infection/care stages in table 2 was adjusted to ensure adequate sample sizes. The number of probable transmission intervals after first viral suppression was too small to enable further

stratification by treatment class. This study was reviewed and approved by the HIV Monitoring Institutional Data Access and Ethics Committee, and reported along STROME-ID guidelines.

Viral sequences of different subtypes (n=355 from MSM), with less than 250 nucleotides (n=368) or indication for intra-subtype recombination (n=52) were removed prior to analysis. Primary drug resistance mutations were masked in each sequence (43). Demographic and clinical data were checked for consistency along patient timelines, and to lie within appropriate ranges. Outliers were reported to the ATHENA quality control team, and manually updated.

### **Recently infected, recipient MSM and infection windows**

We enrolled as recipients all MSM for whom a narrow infection window could be identified. MSM had evidence for infection within 12 months prior to diagnosis if either a last negative HIV-1 antibody test in the 12 months preceding diagnosis, an indeterminate HIV-1 western blot, or clinical diagnosis of acute infection were reported. Figure S1 shows enrollment progress over time. Infection windows were at most 12 months, or shorter if indicated by a last negative HIV antibody test (figure S2).

### **Potential transmitters to recipient MSM**

We enrolled as potential transmitters all registered infected men that overlapped with infection windows of recipients, and thus could have in principle infected a recipient. This definition required estimation of putative infection times. Calculations are based on a method by Rice and colleagues (44), see the online supplementary material. Estimated infection times are associated with substantial uncertainty, and sensitivity analyses were conducted for lower and upper 95% estimates. Table S2 characterizes the potential transmitters to all recipients. Further analysis was restricted to potential transmission pairs with sequences from both individuals (stage B in figure 1).

### **Viral phylogenetic exclusion analysis to construct the transmission cohort**

The viral phylogeny was reconstructed under the GTR nucleotide substitution model with maximum-likelihood methods (45) and is shown in figure S3. 500 bootstrap trees were created to quantify uncertainty in tree reconstruction (14). Genetic distances between sequences from transmitter-recipient pairs were highly variable (figure S4), which was accounted for in all analyses. To guide our choice of exclusion criteria, we considered, first, epidemiologically confirmed transmission pairs from previously published transmission chains in Belgium and Sweden (16, 46). The Belgium transmission chain was subsequently oversampled (15), providing 2,807 sequence pairs from confirmed transmitters and recipients without multi-drug resistance. Further, we considered 4,117 pairs of sequences from the same Dutch patient and 201,605 pairs between Dutch patients that died before the last negative antibody test of another patient. These pairs were used to quantify patterns of viral evolutionary diversification that can be expected among confirmed linked and unlinked pairs, and to develop exclusion criteria with high specificity; see online supplementary material. The Swedish pairs were used for validation purposes. All potential transmitters that



were not excluded were considered phylogenetically probable, and are characterized in table S4.

### **Relative pairwise transmission probabilities**

Among the 2,807 confirmed transmission pairs (15), the genetic distance between sequences from the transmitter and the recipient was strongly associated with the time elapsed between both sampling dates and the midpoint of the established infection window (figure S5). We fitted a probabilistic molecular clock model to these data to describe the relative probability of observing a given genetic distance between sequences from a transmission pair that diverged for a specified amount of time from each other. The fitted model was then used to express the relative probability that a phylogenetically identified transmitter was the actual transmitter to a recipient (figure S5).

### **Matching of clinical data to associate infection/care stages with transmission intervals**

Sources of transmission were not defined in terms of individuals, but the fourteen stages in the infection and care continuum in table 2 (stage E in figure 1). Stages were allocated to transmission intervals based on available clinical data (table S1). The duration of transmission intervals was set to six weeks to accommodate abrupt changes in infection/care stages.

### **Adjusting for censoring and sequence sampling biases**

Towards the present, an increasing fraction of potential transmitters may not have been diagnosed by the time of database closure. Potential transmitters in recent infection at time of diagnosis must, by definition, have been diagnosed within 12 months after the putative transmission interval. Therefore, the extent of right censoring differs between stages. To adjust for right censoring, we counted when potential transmitters in a particular infection/care stage became diagnosed in relation to the time of diagnosis of their recipient (figure S6). This enabled us to estimate the proportion of censored intervals for a hypothetical database closure time in the past (figure S6). We then extrapolated these estimates to the actual database closure time with a bootstrap algorithm; see the online supplementary material. To quantify sequence sampling biases, we compared men with and without a sequence in the near complete population cohort (figure S7). A negative Binomial missing data model was then used to adjust for the number of missing transmission intervals (17). Adjustments accounted for censoring; increasing sampling frequency with duration in care; high sampling frequency of men returning to care, men participating in particular sub-studies, and men with indication of drug-resistance; as well as increasing sampling frequency with calendar time (figure S7).

### **Epidemiological transmission analysis**

Each interval was associated with a phylogenetic transmission probability (stage F in figure 1). The relative pairwise transmission probabilities (figure S5) were equally apportioned to all observed intervals of the same transmitter-recipient pair. Stage-specific data such as viral load was not used to determine these probabilities, to avoid circularity in the attribution of

transmissions to infection/care stages. Then, the transmission probability in an observed interval  $\tau$  from transmitter  $i$  to recipient  $j$  was calculated by

$$p_{ij\tau} = \omega_{ij\tau} / \left( \sum_{k,s} \omega_{kjs} + \sum_z m_j(z) \omega(z) \right),$$

where  $\omega_{ij\tau}$  is the relative transmission probability in interval  $\tau$ , and the denominator sums over all observed, competing intervals as well as expected missing intervals  $m_j(z)$  in stage  $z$  to recipient  $j$ . For missing intervals, relative transmission probabilities were imputed and set to the median  $\omega_{ijs}$  of all observed intervals  $s$  in stage  $z$ , denoted by  $\omega(z)$ . For a missing transmission interval  $v$  in stage  $x$  to recipient  $j$ , we calculated

$$p_{jv} = \omega(x) / \left( \sum_{k,s} \omega_{kjs} + \sum_z m_j(z) \omega(z) \right).$$

In 24 cases, two recipients were each other's phylogenetically probable transmitter. We considered transmission in each direction equally likely. The relative transmission probabilities  $\omega_{ij\tau}$  were calculated by

$$\omega_{ij\tau} = \omega_{ij} \varphi_{ij} / \tau_{ij},$$

where  $\varphi_{ij}$  equals 0.5 if  $i$  and  $j$  are each other's phylogenetically probable transmitters and otherwise one,  $\omega_{ij}$  are the relative pairwise probabilities shown in figure S5, and  $\tau_{ij}$  is the number of transmission intervals between transmitter  $i$  and recipient  $j$ .

These probabilities sum to one per recipient. If all transmitters are sampled, we obtain  $p_{ij\tau} = \omega_{ij\tau} / \sum_{k,s} \omega_{kjs}$ . If some transmitters are not sampled, the first part of the denominator,  $\sum_{k,s} \omega_{kjs}$ , is smaller and adjusted by the second part of the denominator. The number of expected missing intervals  $m_j(z)$  differs by stage, and adjusts for stage-specific censoring and sampling biases.

The proportion of transmissions originating from the fourteen infection/care stages were obtained by summing the corresponding individual-level transmission probabilities (figure S8). Precisely, the proportion of transmissions from stage  $x$  to recipients diagnosed in  $[t_1, t_2]$  was calculated by

$$P^T(x, t_1, t_2) = \frac{\sum_{j \in R(t_1, t_2)} p_j(x)}{\sum_z \sum_{j \in R(t_1, t_2)} p_j(z)} = \frac{1}{J} \sum_{j \in R(t_1, t_2)} p_j(x),$$

where  $R(t_1, t_2)$  is the set of recipients with date of diagnosis in  $[t_1, t_2]$ ,  $J$  is the number of recipients with date of diagnosis in  $[t_1, t_2]$ , and  $p_j(x)$  is the probability that recipient  $j$  was infected by a transmitter in stage  $x$ . The probability  $p_j(x)$  is the sum

$$p_j(x) = \sum_{i \in I_j} \sum_{\tau \in V_{ij}(x)} p_{ij\tau} + \sum_{v=1}^{m_j(x)} p_{jv},$$

where  $I_j$  are the observed, phylogenetically probable transmitters to recipient  $j$ ,  $V_{ij}(x)$  is the set of observed transmission intervals between  $i$  and  $j$  in stage  $x$ , and all other quantities as defined above. The formula for  $P^T(x, t_1, t_2)$  can be intuitively interpreted as the average probability that a recipient was infected by a transmitter in stage  $x$ . Thus, the precision in the estimated  $P^T(x, t_1, t_2)$  depends primarily on the number of available recipients. We identified substantial individual-level variation in the transmission probabilities  $p_j(x)$  (figure S8), suggesting that a relatively large number of past transmission events are needed in order to reliably quantify sources of transmission.

To obtain a central estimate of  $P^T(x, t_1, t_2)$ , we used the central estimates of the  $\omega_{ij\tau}$  and the expected number of missing transmission intervals. To quantify uncertainty in  $P^T(x, t_1, t_2)$ , we propagated uncertainty in the genetic distances and the number of missing transmission intervals with a bootstrap algorithm.

### Epidemiological prevention analysis

With the sources of transmission estimated, we compared the impact of prevention strategies in counterfactual scenarios that modelled the re-distribution of phylogenetically identified transmitters to less infectious stages in the HIV infection and care continuum. This reduced the overall probability that any of the recipients would have been infected to less than one. The proportion of infections that could have been averted in the period  $[t_1, t_2]$  with a counterfactual prevention scenario  $H$  is

$$a(H) = 1 - \sum_{j \in R(t_1, t_2)} \sum_x p_j^H(x),$$

where  $p_j^H(x)$  is the probability that recipient  $j$  is infected by someone in stage  $x$  under the counterfactual prevention scenario  $H$ . The individual-level prevention models are described in the supplementary online material.

### Statistical uncertainty

Central estimates of  $P^T(x, t_1, t_2)$  and  $a(H)$  were obtained under central estimates of the genetic distances in figure S4, the resulting phylogenetic transmission probabilities  $\omega_{ij\tau}$  and the expected number of missing transmission intervals (figure 2C). Bootstrap sampling of the recipients, the empirical distribution of genetic distances, the number of missing transmission intervals under a Negative Binomial missing data model, and the counterfactual re-allocation procedure of probable transmitters to less infectious infection/care stages was conducted to obtain non-parametric 95% confidence intervals. Confidence intervals are based on 1,000 bootstrap replicates.

## Extended acknowledgements

The ATHENA database is maintained by Stichting HIV Monitoring and supported by a grant from the Dutch Ministry of Health, Welfare and Sport through the Centre for Infectious Disease Control of the National Institute for Public Health and the Environment.

### Clinical Centres

\* denotes site coordinating physician

**Academic Medical Centre of the University of Amsterdam:** *HIV treating physicians:* J.M. Prins\*, T.W. Kuijpers, H.J. Scherpbier, J.T.M. van der Meer, F.W.M.N. Wit, M.H. Godfried, P. Reiss, T. van der Poll, F.J.B. Nellen, S.E. Geerlings, M. van Vugt, D. Pajkrt, J.C. Bos, W.J. Wiersinga, M. van der Valk, A. Goorhuis, J.W. Hovius, A.M. Weijnsfeld. *HIV nurse consultants:* J. van Eden, A. Henderiks, A.M.H. van Hes, M. Mutschelknauss, H.E. Nobel, F.J.J. Pijnappel. *HIV clinical virologists/chemists:* S. Jurriaans, N.K.T. Back, H.L. Zaaijer, B. Berkhout, M.T.E. Cornelissen, C.J. Schinkel, X.V. Thomas. **Admiraal De Ruyter Ziekenhuis, Goes:** *HIV treating physicians:* M. van den Berge, A. Stegeman. *HIV nurse consultants:* S. Baas, L. Hage de Looff. *HIV clinical virologists/chemists:* D. Versteeg. **Catharina Ziekenhuis, Eindhoven:** *HIV treating physicians:* M.J.H. Pronk\*, H.S.M. Ammerlaan. *HIV nurse consultants:* E.S. de Munnik. *HIV clinical virologists/chemists:* A.R. Jansz, J. Tjhie, M.C.A. Wegdam, B. Deiman, V. Scharnhorst. **Emma Kinderziekenhuis:** *HIV nurse consultants:* A. van der Plas, A.M. Weijnsfeld. **Erasmus Medisch Centrum, Rotterdam:** *HIV treating physicians:* M.E. van der Ende\*, T.E.M.S. de Vries-Sluijs, E.C.M. van Gorp, C.A.M. Schurink, J.L. Nouwen, A. Verbon, B.J.A. Rijnders, H.I. Bax, M. van der Feltz. *HIV nurse consultants:* N. Bassant, J.E.A. van Beek, M. Vriesde, L.M. van Zonneveld. *Data collection:* A. de Oude-Lubbers, H.J. van den Berg-Cameron, F.B. Bruinsma-Broekman, J. de Groot, M. de Zeeuw- de Man. *HIV clinical virologists/chemists:* C.A.B. Boucher, M.P.G. Koopmans, J.J.A. van Kampen. **Erasmus Medisch Centrum–Sophia, Rotterdam:** *HIV treating physicians:* G.J.A. Driessen, A.M.C. van Rossum. *HIV nurse consultants:* L.C. van der Knaap, E. Visser. **Flevoziekenhuis, Almere:** *HIV treating physicians:* J. Branger\*, A. Rijkeboer-Mes. *HIV nurse consultant and data collection:* C.J.H.M. Duijf-van de Ven. **HagaZiekenhuis, Den Haag:** *HIV treating physicians:* E.F. Schippers\*, C. van Nieuwkoop. *HIV nurse consultants:* J.M. van IJperen, J. Geilings. *Data collection:* G. van der Hut. *HIV clinical virologist/chemist:* P.F.H. Franck. **HIV Focus Centrum (DC Klinieken):** *HIV treating physicians:* A. van Eeden\*. *HIV nurse consultants:* W. Brokking, M. Groot, L.J.M. Elsenburg. *HIV clinical virologists/chemists:* M. Damen, I.S. Kwa. **Isala, Zwolle:** *HIV treating physicians:* P.H.P. Groeneveld\*, J.W. Bouwhuis. *HIV nurse consultants:* J.F. van den Berg, A.G.W. van Hulzen. *Data collection:* G.L. van der Bliet, P.C.J. Bor. *HIV clinical virologists/chemists:* P. Bloembergen, M.J.H.M. Wolfhagen, G.J.H.M. Ruijs. **Leids Universitair Medisch Centrum, Leiden:** *HIV treating physicians:* F.P. Kroon\*, M.G.J. de Boer, M.P. Bauer, H. Jolink, A.M. Vollaard. *HIV nurse consultants:* W. Dorama, N. van Holten. *HIV clinical virologists/chemists:* E.C.J. Claas, E. Wessels. **Maasstad Ziekenhuis, Rotterdam:** *HIV treating physicians:* J.G. den Hollander\*, K. Pogany, A. Roukens. *HIV nurse consultants:* M. Kastelijns, J.V. Smit, E. Smit, D. Struik-

Kalkman, C. Tearno. *Data collection*: M. Bezemer, T. van Niekerk. *HIV clinical virologists/chemists*: O. Pontesilli.

**Maastricht UMC+, Maastricht:** *HIV treating physicians*: S.H. Lowe\*, A.M.L. Oude Lashof, D. Posthouwer. *HIV nurse consultants*: R.P. Ackens, J. Schippers, R. Vergoossen. *Data collection*: B. Weijenberg-Maes. *HIV clinical virologists/chemists*: I.H.M. van Loo, T.R.A. Havenith. **MC Slotervaart, Amsterdam:** *HIV treating physicians*: J.W. Mulder, S.M.E. Vrouenraets, F.N. Lauw. *HIV nurse consultants*: M.C. van Broekhuizen, H. Paap, D.J. Vlasblom. *HIV clinical virologists/chemists*: P.H.M. Smits. **MC Zuiderzee, Lelystad:** *HIV treating physicians*: S. Weijer\*, R. El Moussaoui. *HIV nurse consultant*: A.S. Bosma. **Medisch Centrum Alkmaar:** *HIV treating physicians*: W. Kortmann\*, G. van Twillert\*, J.W.T. Cohen Stuart, B.M.W. Diederer. *HIV nurse consultant and data collection*: D. Pronk, F.A. van Truijen-Oud. *HIV clinical virologists/chemists*: W. A. van der Reijden, R. Jansen. **Medisch Centrum Haaglanden, Den Haag:** *HIV treating physicians*: E.M.S. Leyten\*, L.B.S. Gelinck. *HIV nurse consultants*: A. van Hartingsveld, C. Meerkerk, G.S. Wildenbeest. *HIV clinical virologists/chemists*: J.A.E.M. Mutsaers, C.L. Jansen. **Medisch Centrum Leeuwarden, Leeuwarden:** *HIV treating physicians*: M.G.A. van Vonderen\*, D.P.F. van Houte, L.M. Kampschreur. *HIV nurse consultants*: K. Dijkstra, S. Faber. *HIV clinical virologists/chemists*: J. Weel. **Medisch Spectrum Twente, Enschede:** *HIV treating physicians*: G.J. Kootstra\*, C.E. Delsing. *HIV nurse consultants*: M. van der Burg-van de Plas, H. Heins. *Data collection*: E. Lucas. **OLVG Amsterdam:** *HIV treating physicians*: K. Brinkman\*, G.E.L. van den Berk, W.L. Blok, P.H.J. Frissen, K.D. Lettinga W.E.M. Schouten, J. Veenstra. *HIV nurse consultants*: C.J. Brouwer, G.F. Geerders, K. Hoeksema, M.J. Kleene, I.B. van der Meché, M. Spelbrink, H. Sulman, A.J.M. Toonen, S. Wijnands. *HIV clinical virologists*: M. Damen, D. Kwa. *Data collection*: E. Witte. **Radboudumc, Nijmegen:** *HIV treating physicians*: P.P. Koopmans, M. Keuter, A.J.A.M. van der Ven, H.J.M. ter Hofstede, A.S.M. Dofferhoff, R. van Crevel. *HIV nurse consultants*: M. Albers, M.E.W. Bosch, K.J.T. Grintjes-Huisman, B.J. Zomer. *HIV clinical virologists/chemists*: F.F. Stelma, J. Rahamat-Langendoen. *HIV clinical pharmacology consultant*: D. Burger. **Rijnstate, Arnhem:** *HIV treating physicians*: C. Richter\*, E.H. Gisolf, R.J. Hassing. *HIV nurse consultants*: G. ter Beest, P.H.M. van Bentum, N. Langebeek. *HIV clinical virologists/chemists*: R. Tiemessen, C.M.A. Swanink. **Spaarne Gasthuis, Haarlem:** *HIV treating physicians*: S.F.L. van Lelyveld\*, R. Soetekouw. *HIV nurse consultants*: N. Hulshoff, L.M.M. van der Prijt, J. van der Swaluw. *Data collection*: N. Bermon. *HIV clinical virologists/chemists*: W.A. van der Reijden, R. Jansen, B.L. Herpers, D. Veenendaal. **Stichting Medisch Centrum Jan van Goyen, Amsterdam:** *HIV treating physicians*: D.W.M. Verhagen. *HIV nurse consultants*: M. van Wijk. **St Elisabeth Ziekenhuis, Tilburg:** *HIV treating physicians*: M.E.E. van Kasteren\*, A.E. Brouwer. *HIV nurse consultants and data collection*: B.A.F.M. de Kruijf-van de Wiel, M. Kuipers, R.M.W.J. Santegoets, B. van der Ven. *HIV clinical virologists/chemists*: J.H. Marcelis, A.G.M. Buiting, P.J. Kabel. **Universitair Medisch Centrum Groningen, Groningen:** *HIV treating physicians*: W.F.W. Bierman\*, H. Scholvinck, K.R. Wilting, Y. Stienstra. *HIV nurse consultants*: H. de Groot-de Jonge, P.A. van der Meulen, D.A. de Weerd, J. Ludwig-Roukema. *HIV clinical virologists/chemists*: H.G.M. Niesters, A. Riezebos-Brilman, C.C. van Leer-Buter, M. Knoester. **Universitair Medisch Centrum Utrecht, Utrecht:** *HIV treating physicians*: A.I.M.

Hoepelman\*, T. Mudrikova, P.M. Ellerbroek, J.J. Oosterheert, J.E. Arends, R.E. Barth, M.W.M. Wassenberg, E.M. Schadd. *HIV nurse consultants*: D.H.M. van Elst-Laurijssen, E.E.B. van Oers-Hazelzet, S. Vervoort, *Data collection*: M. van Berkel. *HIV clinical virologists/chemists*: R. Schuurman, F. Verduyn-Lunel, A.M.J. Wensing. **VU medisch centrum, Amsterdam**: *HIV treating physicians*: E.J.G. Peters\*, M.A. van Agtmael, M. Bomers, J. de Vocht. *HIV nurse consultants*: M. Heitmuller, L.M. Laan. *HIV clinical virologists/chemists*: A.M. Pettersson, C.M.J.E. Vandenbroucke-Grauls, C.W. Ang. **Wilhelmina Kinderziekenhuis, UMCU, Utrecht**: *HIV treating physicians*: S.P.M. Geelen, T.F.W. Wolfs, L.J. Bont. *HIV nurse consultants*: N. Nauta.

### Coordinating Centre

*Director*: P. Reiss. *Data analysis*: D.O. Bezemer, A.I. van Sighem, C. Smit, F.W.M.N. Wit. *Data management and quality control*: S. Zaheri, M. Hillebregt, A. de Jong. *Data monitoring*: D. Bergsma, P. Hoekstra, A. de Lang, S. Grivell, A. Jansen, M.J. Rademaker, M. Raethke. *Data collection*: L. de Groot, M. van den Akker, Y. Bakker, M. Broekhoven, E. Claessen, A. El Berkaoui, J. Koops, E. Kruijne, C. Lodewijk, R. Meijering, L. Munjishvili, B. Peeck, C. Ree, R. Regtop, Y. Ruijs, T. Rutkens, L. van de Sande, M. Schoorl, S. Schnörr, E. Tuijn, L. Veenenberg, S. van der Vliet, T. Woudstra. *Patient registration*: B. Tuk.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

We thank the Imperial College High Performance Computing Service (<http://www3.imperial.ac.uk/ict/services/hpc>), three anonymous referees, the HIV treating physicians, HIV nurse consultants and staff of the diagnostic laboratories and facilities in the HIV treatment centres, along with the data collecting and monitoring staff both within and outside the Stichting HIV Monitoring Foundation for their contributions to make this work possible.

**Funding**: OR is supported by the Wellcome Trust (fellowship WR092311MF); CF by the European Research Council (Advanced Grant PBDR-339251) and the Bill & Melinda Gates Foundation (PANGEA-HIV consortium). PR through his institution received independent scientific grant support from Bristol-Myers Squibb, ViiV Healthcare, Gilead Sciences, Janssen Pharmaceuticals Inc., Merck&Co, served on a scientific advisory board for Gilead Sciences and serves on a data safety monitoring committee for Janssen Pharmaceuticals Inc., for which his institution has received remuneration. The Aids Therapy Evaluation in the Netherlands (ATHENA) observational cohort study is part of Stichting HIV Monitoring and supported by a grant from the Netherlands Ministry of Health, Welfare and Sport through its Centre for Infectious Disease Control-National Institute for Public Health and the Environment. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### References

1. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, Hakim JG, Kumwenda J, Grinsztejn B, Pilotto JH, Godbole SV, et al. Fleming. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med*. 2011; 365:493–505. [PubMed: 21767103]
2. Rodger, A.; Bruun, T.; Cambiano, V.; Vernazza, P.; Estrada, V.; Van Lunzen, J.; Collins, S.; Geretti, AM.; Phillips, A.; Lundgren, J. HIV transmission risk through condomless sex if HIV+ partner on suppressive ART: PARTNER Study. 21st, Conference on Retroviruses and Opportunistic Infections; Boston, MA, USA. 2014.
3. Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, Brookmeyer R. Global epidemiology of HIV infection in men who have sex with men. *Lancet*. 2012; 380:367–377. [PubMed: 22819660]



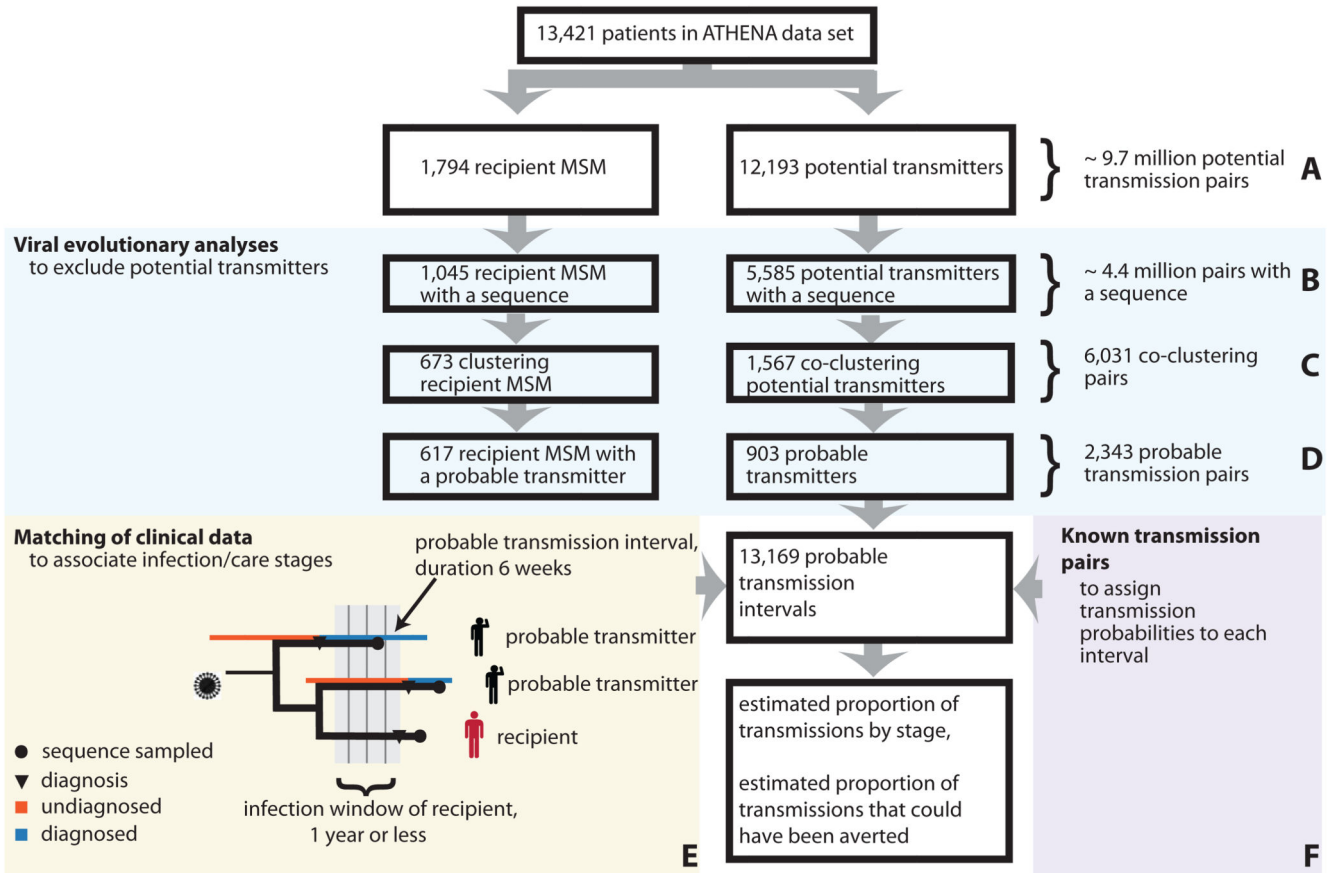
4. Sullivan PS, Carballo-Diequez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, Smith A, Goswami P, Sanchez J. Successes and challenges of HIV prevention in men who have sex with men. *Lancet*. 2012; 380:388–399. [PubMed: 22819659]
5. World Health Organization. No. September 2015. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. Geneva: 2015.
6. Fogarty, A.; Mao, L.; ZM, I.; Santana, H.; Prestage, G.; Rule, J.; Canavan, P.; Murphy, D.; M, D. The Health in Men and Positive Health cohorts: A comparison of trends in the health and sexual behaviour of HIV-negative and HIV-positive gay men, 2002-2005. National Centre in HIV Social Research; Sydney: 2006.
7. Pilcher CD, Fiscus SA, Nguyen TQ, Foust E, Wolf L, Williams D, Ashby R, O'Dowd JO, McPherson JT, Stalzer B, Hightow L, et al. Detection of acute infections during HIV testing in North Carolina. *N Engl J Med*. 2005; 352:1873–1883. [PubMed: 15872202]
8. Weiss HA, Wasserheit JN, Barnabas RV, Hayes RJ, Abu-Raddad LJ. Persisting with prevention: the importance of adherence for HIV prevention. *Emerg Themes Epidemiol*. 2008; 5:8. [PubMed: 18620578]
9. van Sighem, A.; Gras, L.; Kesselring, A.; Smit, C.; Engelhard, I.; Stolte, I.; Reiss, P. Monitoring of human immunodeficiency virus infection in the Netherlands. Amsterdam: 2013. Report 2013
10. Lam TT, Hon CC, Tang JW. Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infections. *Crit Rev Clin Lab Sci*. 2010; 47:5–49. [PubMed: 20367503]
11. Wilson DP. HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. *PLoS Med*. 2012; 9:e1001231. [PubMed: 22807656]
12. Public Health England. Time to test for HIV: Expanding HIV testing in healthcare and community services in England. 2011
13. Romero-Severson E, Skar H, Bulla I, Albert J, Leitner T. Timing and order of transmission events is not directly reflected in a pathogen phylogeny. *Mol Biol Evol*. 2014; 31:2472–2482. [PubMed: 24874208]
14. Pillay D, Rambaut A, Geretti AM, Brown AJ. HIV phylogenetics. *BMJ*. 2007; 335:460–461. [PubMed: 17823148]
15. Vrancken B, Rambaut A, Suchard MA, Drummond A, Baele G, Derdelinckx I, Van Wijngaerden E, Vandamme AM, Van Laethem K, Lemey P. The genealogical population dynamics of HIV-1 in a large transmission chain: bridging within and among host evolutionary rates. *PLoS Comput Biol*. 2014; 10:e1003505. [PubMed: 24699231]
16. Lemey P, Derdelinckx I, Rambaut A, Van Laethem K, Dumont S, Vermeulen S, Van Wijngaerden E, Vandamme AM. Molecular footprint of drug-selective pressure in a human immunodeficiency virus transmission chain. *J Virol*. 2005; 79:11981–11989. [PubMed: 16140774]
17. Little, RJA.; Rubin, DB. Statistical analysis with missing data. Wiley; New York; Chichester: 1987.
18. van Griensven F, Holtz TH, Thienkrua W, Chonwattana W, Wimonsate W, Chaikummao S, Varangrat A, Chemnasiri T, Sukwicha W, Curlin ME, Samandari T, et al. Temporal trends in HIV-1 incidence and risk behaviours in men who have sex with men in Bangkok, Thailand, 2006-13: an observational study. *Lancet HIV*. 2015; 2:e64–70. [PubMed: 26424462]
19. Koedijk FD, van Benthem BH, Vrolings EM, Zuilhof W, van der Sande MA. Increasing sexually transmitted infection rates in young men having sex with men in the Netherlands, 2006-2012. *Emerg Themes Epidemiol*. 2014; 11:12. [PubMed: 25170341]
20. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, Goicochea P, Casapia M, Guanira-Carranza JV, Ramirez-Cardich ME, Montoya-Herrera O, et al. iPrEx Study, Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med*. 2010; 363:2587–2599. [PubMed: 21091279]
21. McCormack S, Dunn DT, Desai M, Dolling DI, Gafos M, Gilson R, Sullivan AK, Clarke A, Reeves I, Schembri G, Mackie N, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. *Lancet*. 2015
22. Molina, JM.; Capitant, C.; Spire, B.; Pialoux, G.; Chidiac, C.; Charreau, I.; Tremblay, C.; Meyer, L.; Delfraissy, JF. CROI; Seattle. 2015.

23. Tanser F, Barnighausen T, Grapsa E, Zaidi J, Newell ML. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*. 2013; 339:966–971. [PubMed: 23430656]
24. Frentz D, Wensing AM, Albert J, Paraskevis D, Abecasis AB, Hamouda O, Jorgensen LB, Kucherer C, Struck D, Schmit JC, Asjo B, et al. Limited cross-border infections in patients newly diagnosed with HIV in Europe. *Retrovirology*. 2013; 10:36. [PubMed: 23551870]
25. Brenner BG, Wainberg MA. Future of phylogeny in HIV prevention. *J Acquir Immune Defic Syndr*. 2013; 63(Suppl 2):S248–254. [PubMed: 23764643]
26. Cope AB, Powers KA, Kuruc JD, Leone PA, Anderson JA, Ping LH, Kincer LP, Swanstrom R, Mobley VL, Foust E, Gay CL, et al. Ongoing HIV Transmission and the HIV Care Continuum in North Carolina. *PLoS One*. 2015; 10:e0127950. [PubMed: 26042804]
27. Skarbinski J, Rosenberg E, Paz-Bailey G, Hall HI, Rose CE, Viall AH, Fagan JL, Lansky A, Mermin JH. Human immunodeficiency virus transmission at each step of the care continuum in the United States. *JAMA Intern Med*. 2015; 175:588–596. [PubMed: 25706928]
28. Rosenberg ES, Millett GA, Sullivan PS, Del Rio C, Curran JW. Understanding the HIV disparities between black and white men who have sex with men in the USA using the HIV care continuum: a modeling study. *Lancet HIV*. 2014; 1:e112–e118. [PubMed: 25530987]
29. I. S. S. Group. Lundgren JD, Babiker AG, Gordin F, Emery S, Grund B, Sharma S, Avihingsanon A, Cooper DA, Fatkenheuer G, Llibre JM, et al. Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *N Engl J Med*. 2015; 373:795–807. [PubMed: 26192873]
30. Volz E, Ionides E, Romero-Severson E, Brandt MG, Mokotoff E, Koopman J. HIV-1 Transmission During Early Infection in Men Who Have Sex with Men: A Phylodynamic Analysis. *PLoS Med*. 2013; 10:e1001568. [PubMed: 24339751]
31. U.S.Fa.D. Administration. Truvada approved to reduce the risk of sexually transmitted HIV in people who are not infected with the virus. 2012
32. Grant RM, Anderson PL, McMahan V, Liu A, Amico KR, Mehrotra M, Hosek S, Mosquera C, Casapia M, Montoya O, Buchbinder S, et al. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men a cohort study. *Lancet Infect Dis*. 2014; 14:820–829. [PubMed: 25065857]
33. Liu A, Cohen S, Follansbee S, Cohan D, Weber S, Sachdev D, Buchbinder S. Early experiences implementing pre-exposure prophylaxis (PrEP) for HIV prevention in San Francisco. *PLoS Med*. 2014; 11:e1001613. [PubMed: 24595035]
34. Bil JP, Davidovich U, van der Veldt WM, Prins M, de Vries HJ, Sonder GJ, Stolte IG. What do Dutch MSM think of preexposure prophylaxis to prevent HIV-infection? A cross-sectional study. *AIDS*. 2015; 29:955–964. [PubMed: 25915169]
35. Mimiaga MJ, White JM, Krakower DS, Biello KB, Mayer KH. Suboptimal awareness and comprehension of published preexposure prophylaxis efficacy results among physicians in Massachusetts. *AIDS Care*. 2014; 26:684–693. [PubMed: 24116985]
36. Mayer KH, Hosek S, Cohen S, Liu A, Pickett J, Warren M, Krakower D, Grant R. Antiretroviral pre-exposure prophylaxis implementation in the United States: a work in progress. *J Int AIDS Soc*. 2015; 18:19980. [PubMed: 26198345]
37. Pao D, Fisher M, Hue S, Dean G, Murphy G, Cane PA, Sabin CA, Pillay D. Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. *AIDS*. 2005; 19:85–90. [PubMed: 15627037]
38. Pant Pai N, Sharma J, Shivkumar S, Pillay S, Vadnais C, Joseph L, Dheda K, Peeling RW. Supervised and unsupervised self-testing for HIV in high- and low-risk populations: a systematic review. *PLoS Med*. 2013; 10:e1001414. [PubMed: 23565066]
39. Lorente N, Preau M, Vernay-Vaisse C, Mora M, Blanche J, Otis J, Passeron A, Le Gall JM, Dhotte P, Carrieri MP, Suzan-Monti M, et al. Expanding access to non-medicalized community-based rapid testing to men who have sex with men: an urgent HIV prevention intervention (the ANRS-DRAG study). *PLoS One*. 2013; 8:e61225. [PubMed: 23613817]
40. Gomez GB, Borquez A, Case KK, Wheelock A, Vassall A, Hankins C. The cost and impact of scaling up pre-exposure prophylaxis for HIV prevention: a systematic review of cost-effectiveness modelling studies. *PLoS Med*. 2013; 10:e1001401. [PubMed: 23554579]

41. Birger RB, Hallett TB, Sinha A, Grenfell BT, Hodder SL. Modeling the impact of interventions along the HIV continuum of care in Newark, New Jersey. *Clin Infect Dis*. 2014; 58:274–284. [PubMed: 24140971]
42. Heuker J, Sonder GJ, Stolte I, Geskus R, van den Hoek A. High HIV incidence among MSM prescribed postexposure prophylaxis, 2000-2009: indications for ongoing sexual risk behaviour. *AIDS*. 2012; 26:505–512. [PubMed: 22156963]
43. Johnson VA, Calvez V, Gunthard HF, Paredes R, Pillay D, Shafer RW, Wensing AM, Richman DD. Update of the drug resistance mutations in HIV-1: March 2013. *Top Antivir Med*. 2013; 21:6–14. [PubMed: 23596273]
44. Rice BD, Elford J, Yin Z, Delpech VC. A new method to assign country of HIV infection among heterosexuals born abroad and diagnosed with HIV. *AIDS*. 2012; 26:1961–1966. [PubMed: 22781226]
45. Kozlov AM, Aberer AJ, Stamatakis A. ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinformatics*. 2015; 31:2577–2579. [PubMed: 25819675]
46. Leitner T, Escanilla D, Franzen C, Uhlen M, Albert J. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. *Proc Natl Acad Sci U S A*. 1996; 93:10864–10869. [PubMed: 8855273]
47. van Sighem A, Nakagawa F, De Angelis D, Quinten C, Bezemer D, de Coul EO, Egger M, de Wolf F, Fraser C, Phillips A. Estimating HIV Incidence, Time to Diagnosis, and the Undiagnosed HIV Epidemic Using Routine Surveillance Data. *Epidemiology*. 2015; 26:653–660. [PubMed: 26214334]
48. Health Protection Agency. Longitudinal analysis of the trajectories of CD4 cell counts. 2011
49. Bezemer D, de Wolf F, Boerlijst MC, van Sighem A, Hollingsworth TD, Fraser C. 27 years of the HIV epidemic amongst men having sex with men in the Netherlands: an in depth mathematical model-based analysis. *Epidemics*. 2010; 2:66–79. [PubMed: 21352777]
50. Eshleman SH, Hudelson SE, Redd AD, Wang L, Debes R, Chen YQ, Martens CA, Ricklefs SM, Selig EJ, Porcella SF, Munshaw S, et al. Analysis of genetic linkage of HIV from couples enrolled in the HIV Prevention Trials Network 052 trial. *J Infect Dis*. 2011; 204:1918–1926. [PubMed: 21990420]
51. Gavryushkina A, Welch D, Stadler T, Drummond AJ. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Comput Biol*. 2014; 10:e1003919. [PubMed: 25474353]
52. McCombs SB, McCray E, Wendell DA, Sweeney PA, Onorato IM. Epidemiology of HIV-1 infection in bisexual women. *J Acquir Immune Defic Syndr*. 1992; 5:850–852. [PubMed: 1517973]
53. Efron, B.; Tibshirani, RJ. An introduction to the bootstrap. Chapman and Hall; Boca Raton: 1998. p. XVIp. 436 Su. a., ed. [Reprint]
54. Cohen MS, Shaw GM, McMichael AJ, Haynes BF. Acute HIV-1 Infection. *N Engl J Med*. 2011; 364:1943–1954. [PubMed: 21591946]
55. Associated Partners of EMIS. EMIS 2010: The European Men-Who-Have-Sex-With-Men Internet Survey. Stockholm: 2013. Findings from 38 countries

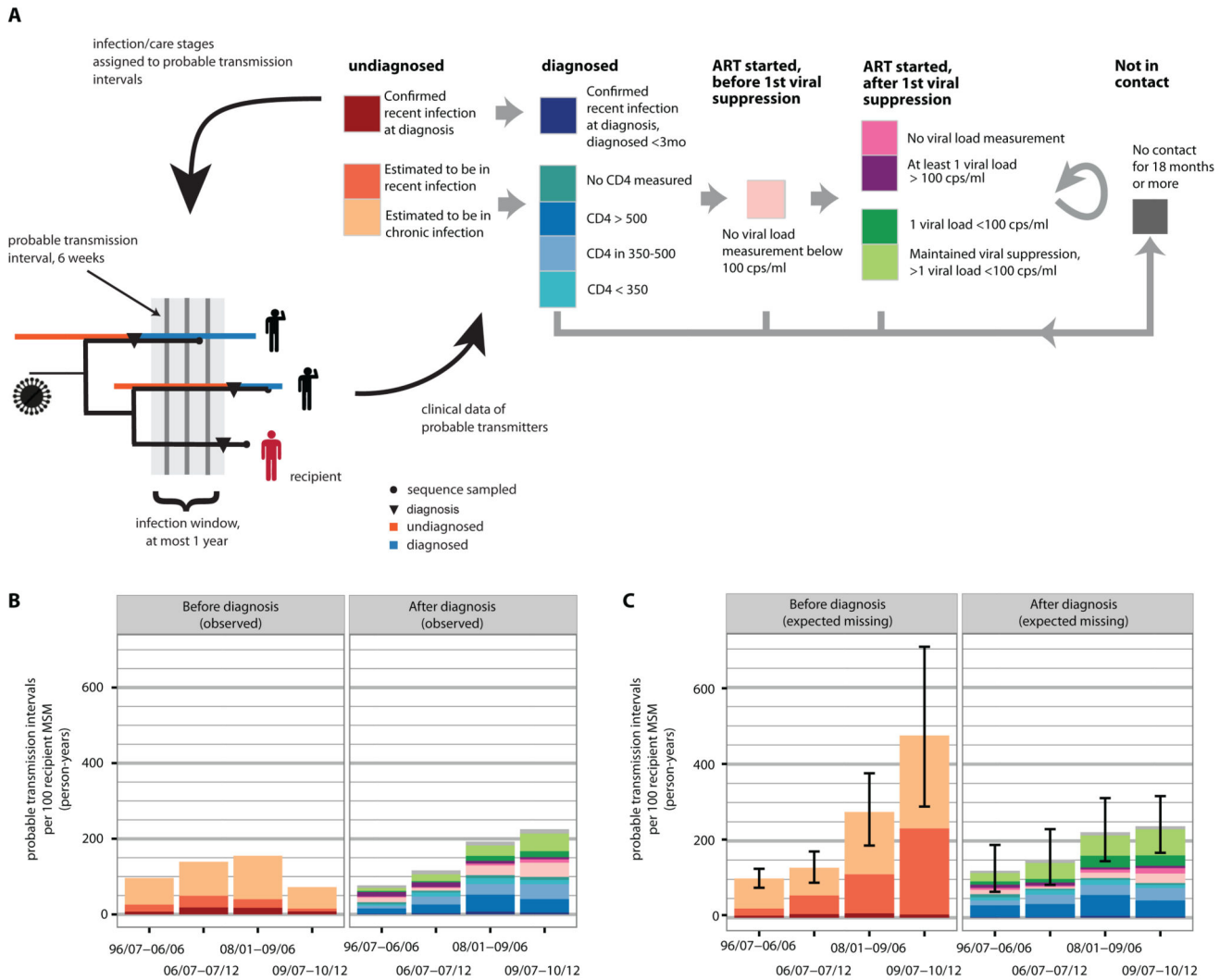
### One Sentence Summary

To tailor HIV prevention strategies amongst men having sex with men, we characterized the sources of ~600 transmission events in the Netherlands. More than half of these infections could have been averted with available antiretrovirals, but only if considerably more men had tested annually.



**Fig. 1. Study design.**

Nationwide sources of transmission were identified for MSM with evidence for recent infection in the first year prior to diagnosis (recipient MSM). (A) Out of all patients in the ATHENA cohort, men whose course of infection overlapped with the infection window were considered as potential transmitters. (B) Only those pairs with sequences from both individuals were considered for further analysis. (C-D) Using viral phylogenetic analyses, the vast majority of pairs could be ruled out. All remaining pairs were considered phylogenetically probable. (E) Based on detailed clinical records, probable transmission events were characterized by stage in the HIV infection and care continuum. Because transmitters progressed in stage over time, we considered time-resolved transmission intervals. (F) Independent viral phylogenetic data from epidemiologically confirmed pairs was used to determine the phylogenetic probability of direct transmission during each interval. Statistical analyses adjusted for extensive sampling and censoring biases.

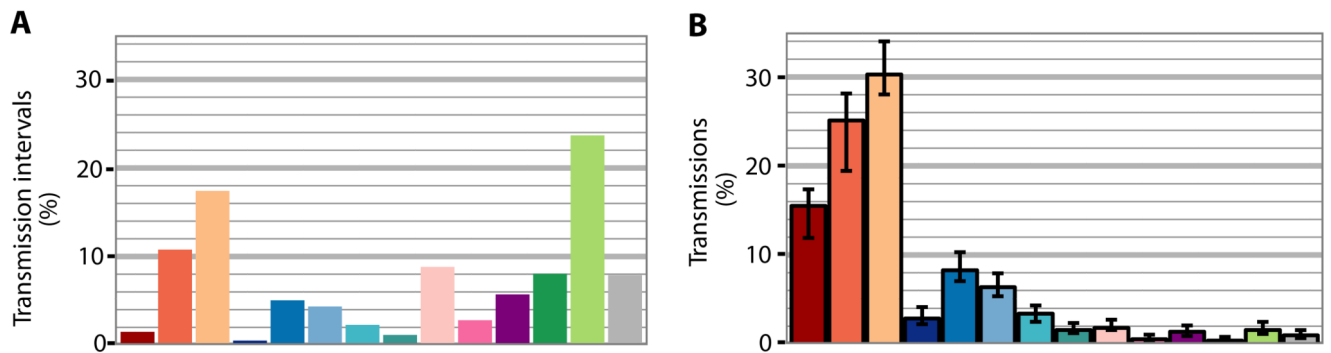


**Fig. 2. Phylogenetically probable transmission intervals, linked to stages in the infection and care continuum.**

(A) Left: Each recipient could have been infected during his infection window from multiple probable transmitters. For each transmitter, the transmission window was split into six-week long probable transmission intervals. Infection/care stages were assigned to these intervals based on clinical data to reflect progression of the transmitters through the infection/care continuum. Right: Relationship between the fourteen infection/care stages as defined in table 2. Transmitters progress uni-directionally, except for stages after first viral suppression, or when individuals re-enter care (as indicated by arrows). (B) For each stage, the total number of observed transmission intervals to recipient MSM during their infection windows is shown. Overall, the number of transmission intervals per recipient increases with time, reflecting the increasing number of infected men in care. Transmitters are increasingly less likely to have been diagnosed by 2013, resulting in a decreasing number of undiagnosed transmission intervals towards the present. (C) In addition to censoring, diagnosed transmitters may not have a sequence sampled. Comparing men with and without a sequence in the near complete population cohort, we could adjust for these biases. The total number of

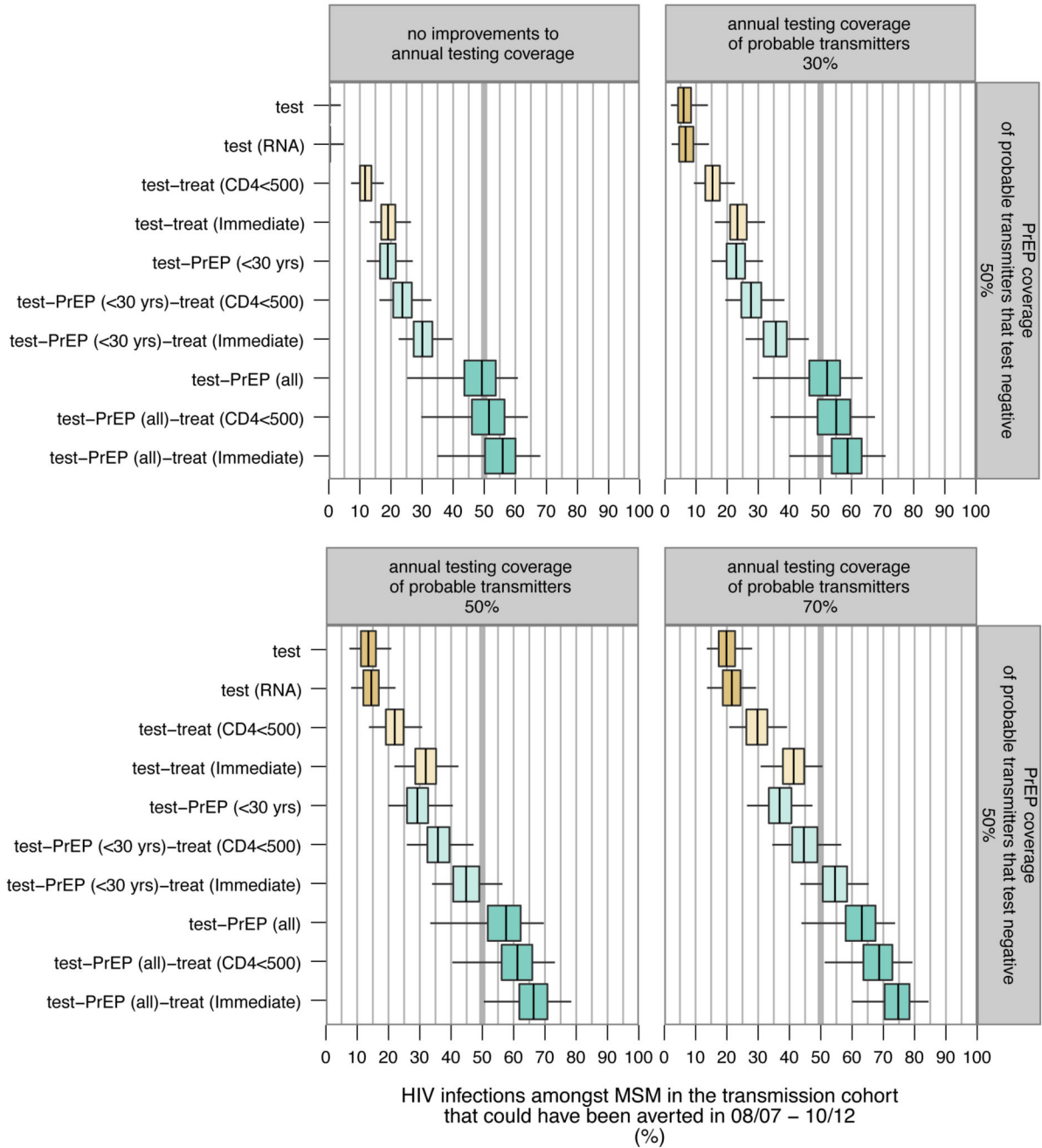


expected missing transmission intervals to recipients diagnosed in one of four observation periods is shown, along with 95% bootstrap confidence intervals. Observed and expected missing transmission intervals were associated with phylogenetic transmission probabilities, which sum to one per recipient.



**Fig. 3. Proportion of transmissions by stage in the infection and care continuum, versus proportion of these stages amongst infected men.**

(A) Relative frequency of infection/care stages in the population, among potential transmitters that overlap with the infection windows of recipient MSM and could have in principle transmitted to one of the recipient MSM. (stage A in figure 1, colour codes as in figure 2). (B) Proportion of the 617 transmission events attributable to each infection/care stage (bar: 95% bootstrap confidence interval).



**Fig. 4. Impact of biomedical interventions amongst MSM in the Netherlands.**

Estimated proportion of transmissions that could have been averted in the period 2008/07-2010/12 if the corresponding additional prevention strategies had been implemented by 2008/07 (line: median, box: bootstrap interquartile range, whiskers: 95% bootstrap confidence interval). Scenarios were varied by annual testing coverage of phylogenetically identified, probable transmitters. Current testing coverage was 17%,

corresponding to the proportion of probable transmitters that had a negative test in the twelve months prior to diagnosis.

**Table 1**  
**HIV incidence trends and care for infected MSM in the Netherlands and other countries.**

Country	Annual testing of uninfected MSM		Diagnosed MSM receiving ART			Treated MSM with suppressed viral load			Retention of MSM in care			HIV incidence among MSM		
	Year	%	Year	%	Median CD4 count at ART initiation (cells/ml)	Year	%	Viral load threshold (cps/ml)	Year	%	Year	%	Year	Trend
Netherlands	2003	??	2003	79	202	2003	80	<100	2003	92	2003	92	2003	Increasing <sup>n</sup>
	2013	38.4 <sup>a</sup>	2013	90	382	2013	91	<100	2012	95	2013	95	2013	Stable <sup>n</sup>
Australia	2013	61.1 <sup>b</sup>	2013	75 <sup>b</sup>	379 <sup>b</sup>	2013	88 <sup>k</sup>	<50	2013	96 <sup>h</sup>	2013	96 <sup>h</sup>	2013	Increasing, stable in Western Australia and Queensland <sup>o</sup>
British Columbia	2009	51 <sup>c</sup>	2014	85 <sup>e</sup>	411 <sup>e</sup>	2014	84 <sup>e</sup>	<50	2011	86 <sup>d</sup>	2013	86 <sup>d</sup>	2013	Stable <sup>p</sup>
Switzerland	2010	39.3 <sup>a</sup>	2014	86 <sup>f</sup>	402 <sup>f</sup>	2012	96 <sup>l</sup>	<200	2012	97 <sup>i</sup>	2014	97 <sup>i</sup>	2014	Decreasing new diagnoses and recent infections <sup>q</sup>
United Kingdom	2010	36.4 <sup>a</sup>	2013	86 <sup>g</sup>	420 <sup>j</sup>	2013	91 <sup>m</sup>	<200	2013	95 <sup>m</sup>	2013	95 <sup>m</sup>	2013	Stable <sup>r</sup>

<sup>a</sup>From The EMIS Network. EMIS 2010: The European Men-Who-Have-Sex-With-Men Internet Survey. Findings from 38 countries, Stockholm: European Centre for Disease Prevention and Control, 2013.

<sup>b</sup>From Gay Community Periodic Surveys. <https://kirby.unsw.edu.au/projects/gay-community-periodic-surveys>, reported in HIV, hepatitis, and sexually transmissible infections in Australia. Annual surveillance report 2014.

<sup>c</sup>From Mancount, prospective cross-sectional survey in Vancouver [http://www.mancount.ca/files/Mancount\\_Report2010.pdf](http://www.mancount.ca/files/Mancount_Report2010.pdf).

<sup>d</sup>From Nosyk B, Montaner JS, Colley G, et al. The cascade of HIV care in British Columbia, Canada, 1996-2011: a population-based retrospective cohort study. The Lancet infectious diseases 2014;14:40-9

<sup>e</sup>From HIV monitoring quarterly report for British Columbia, Fourth quarter 2014.

<sup>f</sup>2621 out of 3081 MSM on ART and registered in the Swiss HIV Cohort Study, personal communication with the Datacenter of the Swiss HIV Cohort Study.

<sup>g</sup>From <https://www.gov.uk/government/statistics/hiv-data-tables>.

<sup>h</sup>From Australian HIV Observational Database Annual Report 2014, reporting care indicators in a closed observational cohort.

<sup>i</sup>From <http://www.shcs.ch/155-shcs-key-data-figures> update June 2014.

<sup>j</sup>Within 9 months prior to ART initiation, personal communication PHE.

<sup>k</sup>From the Australian HIV Observational Database, reported in HIV, hepatitis, and sexually transmissible infections in Australia. Annual surveillance report 2014.

<sup>l</sup>Köhler P, Schmidt AJ, Ledergerber B, Vernazza P. CROI2015. <http://www.croiconference.org/sites/default/files/posters-2015/1008.pdf>.

<sup>77</sup>From HIV in the United Kingdom: 2014 Report.

<sup>78</sup>From reference (44).

<sup>9</sup>From fact sheet HIV and AIDS in Australia, 20th International AIDS conference.

<sup>8</sup>From <http://www.phac-aspc.gc.ca/aids-sida/publication/epi/2010/index-eng.php>

<sup>9</sup>From HIV- und STI-Fallzahlen 2014: Berichterstattung, Analysen und Trends, in comparison with numbers for 2008 in the 2012 report, [http://www.bag.admin.ch/hiv\\_aids/12472/12480/12481/12484/index.html?lang=de](http://www.bag.admin.ch/hiv_aids/12472/12480/12481/12484/index.html?lang=de).

<sup>7</sup>From Birrell PJ, Gill ON, Delpech VC, et al. HIV incidence in men who have sex with men in England and Wales 2001-10: a nationwide population study. *The Lancet infectious diseases* 2013;13:313-8.

<sup>\*</sup> Estimate not specific to MSM.



**Table 2**  
**Stages in the HIV infection and care continuum**

<b>Infection/care stage of transmitter</b>	<b>Definition</b>
<b>Undiagnosed</b>	Transmission intervals whose midpoint is before diagnosis:
Confirmed recent infection at diagnosis	All transmission intervals of transmitters that were in laboratory confirmed recent infection at time of diagnosis.
Estimated to be in recent infection	Considering transmitters that had no evidence for recent infection at time of diagnosis, all transmission intervals whose midpoint is less than 12 months after the estimated infection date.
Estimated to be in chronic infection	Considering transmitters that had no evidence for recent infection at time of diagnosis, all transmission intervals whose midpoint is more than 12 months after the estimated infection date.
<b>Diagnosed</b>	Transmission intervals whose midpoint is after diagnosis and before ART start (only of transmitters that are in contact with care services):
Diagnosed < 3mo, Recent infection at diagnosis	Considering potential or probable transmitters that were in laboratory confirmed recent infection at time of diagnosis, all transmission intervals whose midpoint is within the first three months after diagnosis.
No CD4 measured	No available CD4 count since diagnosis up to the midpoint of the interval.
CD4 > 500	CD4 counts remained above 500 cells/ml between the first CD4 count up to the midpoint of the interval.
CD4 in [350-500]	CD4 counts decreased to 350-500 cells/ml between the first CD4 count up to the midpoint of the interval.
CD4 < 350	CD4 counts decreased to below 350 cells/ml between the first CD4 count up to the midpoint of the interval.
<b>ART initiated</b>	Transmission intervals whose midpoint is after ART start (only of transmitters that are in contact with care services):
Before first viral suppression	No first viral load measurement below 100 copies/ml in any transmission interval of the transmitter after ART start
After first viral suppression <sup>¶</sup>	
No viral load measured <sup>¶</sup>	No viral load measurement in any transmission interval of the transmitter after ART start
No viral suppression <sup>¶</sup>	At least one viral load measurement at or above 100 copies/ml in any transmission interval of the transmitter after ART start
Viral suppression, one observation <sup>¶</sup>	One viral load measurement in any transmission interval of the transmitter after ART start, which is below 100 copies/ml.
Viral suppression, >1 observations <sup>¶</sup>	Several viral load measurements in any transmission interval of the transmitter after ART start, all of which are below 100 copies/ml.
<b>Not in contact</b>	No patient record (last contact, clinic visit, CD4 measurement, viral load measurement) in the past and future 9 months from the midpoint of the transmission interval.

<sup>¶</sup>While flow through the stages is typically unidirectional, men could move freely between these stages.

**Table 3**  
**Characteristics of the recipient MSM with identified sources of transmission**

Characteristic	Recipient MSM with a phylogenetically probable transmitter (n= 617)	Recipient MSM with or without a sequence (n= 1,794)	Diagnosed MSM (n= 7,978)
<b>Evidence for infection in the past year</b>			
<b>Previous negative test in the past year (%)</b>	77	76	17
<b>Laboratory diagnosis (%)</b>	8	7	2
<b>Clinical diagnosis of acute infection (%)</b>	15	17	4
<b>Age at diagnosis (years; mean and IQR)</b>	36.8 (29.5-42.9)	37.2 (29.9-43.5)	38.7 (31.3-45.1)
<b>First CD4 count within 12 months of diagnosis and before ART start (cells/ml; mean and IQR)</b>	505 (350-630)	534 (360-670)	402 (200-560)
<b>Viral load count within 12 months of diagnosis (log<sub>10</sub> RNA; mean and IQR)</b>	4.9 (4.4-5.5)	4.8 (4.3-5.4)	4.7 (4.3-5.3)
<b>In care in the Amsterdam metropolitan area (%)</b>	45.1	43.5	43.6
<b>Last negative test within 12 months prior to diagnosis (%)</b>	77.0	76.1	17.1
<b>Self-reported in country infection (%)<sup>¶</sup></b>	96.9	91.9	88.5

<sup>¶</sup>Of those self-reporting a country of origin.

**Table 4**  
**Proportion of transmissions by stage in the HIV infection and care continuum.**

Infection/care stage of transmitter	% of transmissions by time of diagnosis of recipient MSM (95% confidence interval)				
	Overall (n=617)	96/07-06/04 (n=165)	06/05-07/12 (n=145)	08/01-09/06 (n=151)	09/07-10/12 (n=156)
<b>Undiagnosed (total)</b>	70.9 (65.8-72.5)	67.6 (59.3-72.7)	72.3 (64.2-76.9)	71.8 (63.4-76.3)	72.2 (63.3-76.3)
Confirmed recent infection at diagnosis	15.5 (11.9-17.4)	15 (7.6-19.4)	21.7 (15-26.5)	16.4 (11-20.8)	9.4 (5.6-14.1)
Estimated to be in recent infection	25.1 (19.4-28.1)	17.3 (11.7-22.7)	23 (15.1-30.1)	25.9 (15.4-33.6)	34.6 (19.4-43.4)
Estimated to be in chronic infection	30.3 (28-34)	35.2 (30.2-42)	27.6 (22.4-34)	29.5 (24.2-36.1)	28.2 (23-35.7)
<b>Diagnosed (total)</b>	22.4 (20.7-26.2)	23.6 (18.5-29.7)	22.9 (18.6-29.1)	22.8 (18.3-29.4)	20.7 (17.4-27.3)
Diagnosed < 3mo, Recent infection at diagnosis	2.9 (2.2-4.1)	2.5 (1-4.9)	3.2 (1.7-5.5)	3 (1.9-5.4)	2.8 (1.8-4.4)
No CD4 measured	1.6 (1.2-2.4)	2.9 (1.6-4.8)	0.8 (0.4-1.8)	1.5 (0.6-3)	1 (0.6-2.1)
CD4 > 500	8.3 (7-10.3)	10.2 (6.7-14.2)	7 (4.5-10.8)	8.7 (5.9-12.5)	7.1 (5.4-10.1)
CD4 in [350-500]	6.4 (5.4-7.9)	4.8 (2.6-7.8)	7.3 (5.1-10.5)	5.9 (4.2-8.3)	7.7 (5.7-11)
CD4 < 350	3.4 (2.5-4.3)	3.2 (1.2-5.5)	4.6 (2.6-6.6)	3.7 (2.2-5.6)	2.1 (1.3-3.3)
<b>ART initiated (total)</b>	5.7 (5.2-7.8)	7 (4.8-11.7)	3.7 (2.2-6.5)	4.9 (3.7-8.1)	6.7 (5.4-10.2)
Before first viral suppression	1.8 (1.6-2.7)	2.2 (1.2-4.4)	0.7 (0.4-1.5)	1.3 (0.9-2.6)	2.8 (2.1-4.6)
After first viral suppression					
No viral load measured	0.5 (0.3-1)	0.9 (0.1-2.4)	0.1 (0-0.3)	0.3 (0.1-0.9)	0.8 (0.4-1.8)
No viral suppression	1.4 (0.9-2.1)	2.8 (1.2-5.2)	1.2 (0.4-2.6)	0.9 (0.4-1.9)	0.5 (0.1-1)
Viral suppression, one observation	0.4 (0.3-0.8)	0.1 (0-0.5)	0.2 (0-0.8)	0.5 (0.2-1.7)	0.6 (0.3-1.5)
Viral suppression, >1 observations	1.6 (1.1-2.5)	1 (0.1-2.6)	1.5 (0.6-3.1)	1.9 (0.9-3.6)	2 (1.1-3.6)
<b>Not in contact</b>	1 (0.7-1.6)	1.8 (0.8-3.4)	1.1 (0.4-2.3)	0.5 (0.2-1.4)	0.4 (0.2-0.8)
<b>Recent infection (total)</b>	43.5 (36.6-46)	34.9 (25.4-40.6)	47.9 (36.9-54.8)	45.3 (33.3-54.1)	47.7 (32.8-53.8)