Impact of antiretroviral therapy on HIV-1 transmission dynamics

Bezemer, D.O.

Citation for published version (APA):
A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy

Daniela Bezemer\textsuperscript{a}, Frank de Wolf\textsuperscript{a,b}, Maarten C. Boerlijst\textsuperscript{c}, Ard van Sighem\textsuperscript{a}, T. Deirdre Hollingsworth\textsuperscript{b}, Maria Prins\textsuperscript{d,e}, Ronald B. Geskus\textsuperscript{d,f}, Luuk Gras\textsuperscript{a}, Roel A. Coutinho\textsuperscript{g,h} and Christophe Fraser\textsuperscript{b}

Objective: Reducing viral load, highly active antiretroviral therapy has the potential to limit onwards transmission of HIV-1 and thus help contain epidemic spread. However, increases in risk behaviour and resurgent epidemics have been widely reported post-highly active antiretroviral therapy. The aim of this study was to quantify the impact that highly active antiretroviral therapy had on the epidemic.

Design: We focus on the HIV-1 epidemic among men who have sex with men in the Netherlands, which has been well documented over the past 20 years within several long-standing national surveillance programs.

Methods: We used a mathematical model including highly active antiretroviral therapy use and estimated the changes in risk behaviour and diagnosis rate needed to explain annual data on HIV and AIDS diagnoses.

Results: We show that the reproduction number $R(t)$, a measure of the state of the epidemic, declined early on from initial values above two and was maintained below one from 1985 to 2000. Since 1996, when highly active antiretroviral therapy became widely used, the risk behaviour rate has increased 66%, resulting in an increase of $R(t)$ to 1.04 in the latest period 2000–2004 (95% confidence interval 0.98–1.09) near or just above the threshold for a self-sustaining epidemic. Hypothetical scenario analysis shows that the epidemiological benefits of highly active antiretroviral therapy and earlier diagnosis on incidence have been entirely offset by increases in the risk behaviour rate.

Conclusion: We provide the first detailed quantitative analysis of the HIV epidemic in a well defined population and find a resurgent epidemic in the era of highly active antiretroviral therapy, most likely predominantly caused by increasing sexual risk behaviour.

Keywords: antiretroviral therapy, homosexual men, infectious diseases, mathematical models, models/projections, sexual behaviour, surveillance

Introduction

To determine the success of a decade of widespread intervention with highly active antiretroviral therapy (HAART) on controlling the human immunodeficiency virus type 1 (HIV-1) epidemic, we analysed the transmission dynamics of HIV-1 over the past 25 years among men having sex with men (MSM) in the Netherlands.

The first AIDS cases in the Netherlands among MSM were identified in 1981 [1]. HAART, consisting of a combination of drugs from three, and later four, different drug classes, became widely available from 1996 onwards. HAART dramatically reduced plasma and seminal viral load [2,3], resistance [4] and mortality rates [5,6]. As infectivity is shown to be strongly correlated to viral load [7,8], the widespread use of HAART might thus be expected to have reduced the incidence of HIV infections. Paradoxically, resurgent epidemics have been widely reported post-HAART [9,10]. Increases in risk behaviour [11–13] and syphilis and gonorrhoea diagnoses have also been documented in populations of MSM in several developed countries [13–15]. Earlier mathematical modelling studies have demonstrated that an increase in risk behaviour has the potential to counterbalance the beneficial effect of HAART [16–23].

In the present study, we aim to evaluate the separate impact of risk behaviour, HIV testing behaviour and HAART on the HIV epidemic in Dutch MSM by means of a mathematical model fitted to data recorded within several national databases, which provided extensive information on epidemic trends.

Methods

Model

A mathematical model describing HIV transmission and HAART use among MSM in the Netherlands was constructed. The model we used described natural (untreated) disease progression, diagnosis and subsequent use of HAART. The basic structure of the model is illustrated in Fig. 1. The modelling strategy was tailored to the task of analyzing annual HIV and AIDS diagnosis time series, and specifically to tracking changes in per capita transmission rates. The most important factor in this respect is a simultaneous estimation of the prevalence of infectious individuals, weighted by their relative infectiousness, which depends on stage of infection and treatment status, and the incidence of new infections. Mathematical details and analyses of the model, including sensitivity analyses, hypothetical scenarios and predictions, and further data are available on request from the authors.

Survival distribution

A method to increase realism in compartmental models is to include a unidirectional flow through several compartments, corresponding to an Erlang survival distribution. By fitting such a distribution to data from 130 MSM seroconverters before the HAART era in the Amsterdam Cohort Studies [24], the maximum likelihood estimate corresponded to five compartments with mean stay in each of 1.89 years. Patients starting their disease progression first spend on average 0.24 year in an extra initial compartment that represents primary infection, and we equated the last stage of infection with AIDS, an approximation that seemed reasonable given the match with the estimated duration of high transmissibility that

![Fig. 1. Model structure. (a) Flow diagram of model of HIV-1 transmission among MSM. The model describes progression through different stages of natural history and treatment. Arrows depict the different flow rates between compartments. New infections start with primary HIV infection and then progress to death through stages of infection. Undiagnosed infections get diagnosed, after which risk behaviour can be reduced. The diagnosis rate varies over time. From 1996, antiretroviral treatment is available that can be long-term or transiently successful. Disease progression is represented by stacks: white stacks represent stages that are infectious and in which disease progression occurs. The nature of disease progression within a stack is shown in detail in (b). Grey stacks represent stages that are not infectious and in which disease progression does not occur as viral replication is suppressed by treatment. Infectiousness is highest during primary infection and AIDS, and lower during stages 1–4. All stages, weighted by their relative infectiousness and fitted by the risk behaviour rate parameter, contribute to the estimated annual new infections such that the annual data on HIV and AIDS diagnosis can be described. Imported infections flow into primary infection and undiagnosed compartments. (b) To enhance realism on survival distribution in the model, disease progression is represented by a unidirectional flow through five compartments with mean stay in each of 1.89 years.](image-url)
has been observed prior to death [25]. The relative infectiousness of each stage is calculated from Wawer et al. [25], with primary and last stages of infection being more infectious than the other four stages. As 97% of the MSM population is infected with subtype B, imported infections are all assumed to be in the primary infection stage entering the Netherlands after short holidays, but this assumption was not critical (unpublished analysis available on request).

**Highly active antiretroviral therapy**
In the HAART era, patients start HAART after being diagnosed, and they will be either on successful treatment (no detectable viral load) or experience treatment failure, with a viral rebound and infectiousness [26]. During successful treatment, HAART is assumed to block both HIV transmission and disease progression [2–8,27,28], and viral blips are not taken into account [29]. Patients experiencing treatment failure are assumed to have periods of apparent successful treatment before failure [26]. After failure, patients go through the unidirectional stages of natural disease progression. We assume that there are three HAART treatment opportunities before patients fail completely and progress to death (representing the diversity of treatments available) [26]. The HAART era started in 1995 with clinical trials and compassionate use, and the mass treatment programme started in 1996 and was fully implemented by 1998. The influence of pre-HAART therapy on HIV viral load and transmission before 1995 is neglected. People start HAART with a rate irrespective of their stage of infection (as multistage disease progression is included in the model, this assumption approximately reproduces the observed pattern of HAART initiation), but at the AIDS stage people are set on HAART immediately. Disease progression is unidirectional. Parameters on HAART use and failure were obtained from the ATHENA national observational cohort [26].

**Transmission rate and risk behaviour**
The standardized per infectious capita transmission rate $\beta(t)$ is a time-varying function that measures the relative rate at which an HIV-positive infectious individual infects new individuals. It is standardized by setting it equal to 1.0 for untreated, undiagnosed individuals in the asymptomatic stage of infection during the first phase of the epidemic (1980–1983), so that all other values are measured relative to this. It is primarily intended as a measure of changes in risk behaviour that can be estimated in our study, and for convenience $\beta(t)$ will be referred to as risk behaviour rate. $\beta(t)$ is in fact a compound measure that is affected by changes in the partner change rate, by the rate and nature of risky sex acts within partnerships, by the effect of 'saturation' of the susceptible population (when new sexual partners are already previously infected) and by the effect of the changing prevalence of other sexually transmitted infections (STIs) in modulating HIV transmission.

Parameters in the model explicitly adjust for the effect of HAART in reducing infectiousness, for the increased infectiousness during primary and late (AIDS) stages of disease and for the effect of diagnosis in reducing risk behaviour. We assumed that MSM have a 50% reduction in risk behaviour after becoming aware of their seropositive status and implemented this into our model [30]. These assumptions were all encoded as disease-stage-specific scaling parameters of risk behaviour rate $\beta(t)$.

**Reproduction number**
We define the reproduction number $R(t)$ as the average number of people an infected person at time $t$ would infect over his whole infectious lifespan if conditions remained the same as at time $t$ [31,32]. It incorporates all factors including risk behaviour, effect of diagnosis and the effects of treatment with HAART in preventing infection. If the within-country $R(t)$ is greater than 1, then the epidemic will grow exponentially driven by local transmission, and conversely if this number is less than 1, the epidemic will contract down to a number proportional to the number of imported cases [31]. It is a key aim of public health interventions to avoid a locally driven epidemic and maintain $R(t)$ below one. The state of the epidemic can be characterized by $R(t)$ that can be calculated from the best fit parameters in the model.

**Model fit**
We fitted our model simultaneously to the observed time series of annual new diagnoses [32] and annual new AIDS cases (see below) [14,32,33], which are constrained by the diagnosis rate and the risk behaviour rate $\beta(t)$, and this made it feasible to estimate both these unknown parameters. Changing these independent parameters has different effects, which differently affect the goodness of fit of the model to the time series. Increasing risk behaviour increases both the number of diagnoses and AIDS cases, whereas increasing the diagnosis rate increases the number of diagnoses in the short term but leads to sustained long-term reductions in the number of diagnoses and AIDS cases.

The analysis was stratified into four distinct historical intervals: 1980–1983, the first AIDS cases were diagnosed [1]; 1984–1995, serological testing became available, increasing HIV awareness, introduction of first mono- and dual-antiretroviral and dual-antiretroviral therapies [3,6]; 1996–1999, early HAART era; and 2000–2004, current HAART era. The diagnosis rate during asymptomatic stages was estimated but was assumed to be zero during the first period (1980–1983). Diagnosis was assumed to be rapid (within 1 month) after AIDS, whereas zero during primary infection. The mean time to diagnosis, defined under conditions at time $t$, was calculated from the estimated diagnosis rate. The epidemic is assumed to have started with an import of cases in 1980. The model was solved numerically using Runge-Kutta 4 algorithm and
was fitted to data by a custom maximum likelihood method. All analyses were performed in Berkeley Madonna, version 8.0.1 (http://www.berkeleymadonna.com).

Annual new AIDS cases
To account for the effect of HAART in preventing progression to AIDS, we used different data sets to simultaneously fit to the following: before 1997, the model is fitted to annual data on AIDS diagnosis among MSM and collected by Statistics Netherlands from the beginning of the HIV epidemic [14,34]; from 1996, the model is fitted to annual data of number of MSM getting HIV diagnosed while having AIDS in the ATHENA national observational cohort [33].

Annual new diagnoses
From 1984 the model was fitted to data on annual diagnoses per year among MSM in the ATHENA national observational cohort [33]. Since 1998 all HIV patients in the Netherlands have been registered and monitored as part of the ATHENA national observational cohort. The year of first HIV diagnosis is recorded retrospectively at the point of registration into ATHENA. Patients who received HAART and died in the period 1996–1997 were included in the ATHENA database retrospectively. Although there is some uncertainty on the completeness of the retrospective inclusion, it is expected to have only minor bearing on our results. MSM who died before 1996 are not included in the ATHENA database. We explicitly accounted for data truncation process in our model by implement chance of surviving until 1996 for the respective stage infection. In parallel, a prediction is made of the true (truncated) curve of the number of new diagnoses (Fig. Data from 2005 are still incomplete, and they are thus included in the current study.

Source of infection
By the start of 2005, 3516 MSM diagnosed with HIV been included in the ATHENA observational cohort these registered infections, 8% were reported to have acquired the infection while abroad and 62% from a partner within the Netherlands. Of those born in Netherlands, 4% were infected abroad and of those abroad, 41% were infected abroad. We assume that remaining 30% of infections with an unknown country of infection are split according to these respective ratios respective to their country of birth. Thus, we estim that overall 14% of diagnosed infections are importe For model verification, we compared the model number of prevalent cases with number of living HIV-positive MSM in the ATHENA database. Also, data on predicted annual number of deaths with documented annual AIDS deaths in the Netherlands [14] were used for model outcome verification. These data con

Fig. 2. Data and model fit. (a) Number of new diagnoses of HIV. Thick lines and dots, cases acquired within the Netherlands; lines and triangles, cases acquired abroad. Empty symbols represent years when data are only available for patients surviving 1996, and dashed lines represent estimated actual number of diagnoses. (b) Number of new diagnoses of AIDS. Data from Di Health Inspectorate (black dots) used in model fit, and ATHENA (empty dots) for model verification (not fitted). (c) Estimate the risk behaviour rate $b(t)$ (solid line, left axis; 1.30, 0.56, 0.66, 0.93) and the mean time between infection and diagnosis (dashed line, right axis; 7.88, 3.71, 3.16, 2.90). (d) Estimate of the reproduction number $R(t)$ (solid line, left axis; 2.39, 0.89, 0.76, 1.04) of the number of new infections acquired within Netherlands (dashed line, right).
After 1995, when HAART was introduced, the reproduction number declined yet further to 0.76 (95% CI 0.7–0.86), but the reduction was not as great as it could have been due to a 18% (95% CI 3–34%) increase of the risk behaviour rate, \( \beta(t) \). The risk behaviour rate is estimated to have increased yet further over the period 2000–2004 and returned to only 29% (95% CI 22–72%) below its value in the initial period 1980–1983. Reductions in the estimated mean time from infection to diagnosis [from 3.71 years (95% CI 3.49–3.97) in 1984–1995 to 2.90 years (95% CI 2.84–3.03) in 2000–2004] with consecutive reductions in risk behaviour and widespread treatment with HAART resulted in the reproduction number being much lower than in the initial time period 1980–1983. Still, \( R(t) \) for the last time period 2000–2004 is estimated to be 1.04 (95% CI 0.98–1.09), near or above the critical epidemic threshold, and thus indicating that HIV may once again be spreading epidemically among MSM in the Netherlands.

From the best fit model, we estimated that 24% of all living HIV-positive MSM were unaware of their HIV-positive status at the start of 2005 and that they account for 90% of new infections. Without both the increase of the risk behaviour rate and the decrease of time to diagnosis, the reproduction number \( R(t) \) would have decreased by 24% from 0.89 to 0.68 due to the introduction of HAART. The risk behaviour rate would need to increase by 32% to offset this benefit, with 43% in order to offset the simultaneous benefits of the increase in testing behaviour and with 59% in order to get the reproduction number \( R(t) \) equal to one, that is, to revert to epidemic growth. An increase of 66% was measured to have occurred. On the basis of these model estimates, we conclude that HAART has played an important role in limiting transmission but that any gains made have been more than offset by increases in the risk behaviour rate. Had these increases not occurred in the HAART era, the reproduction number \( R(t) \) would have declined to 0.6, and the epidemic would have been in convinced decline.

We verified our predictions subjectively for consistency with approximated data on annual number of AIDS deaths in MSM (see Methods), and on the number of currently living diagnosed individuals in the national patient database ATHENA [33], shown in Fig. 3, and on the number of annual AIDS diagnoses after 1996 (Fig. 2b). We considered the quality of fit acceptable given that the model was not fitted to these data. A qualitative comparison of CD4 cell counts at diagnosis with model predictions in terms of disease stages shows similar trends.

Fig. 3. Consistency of model fit. (a) Number of deaths caused by HIV. Seventy percent of number of AIDS deaths among male (black dots, see methods) and model prediction of AIDS deaths among MSM (thick line). Deaths among MSM in ATHENA (empty dots). (b) Number of prevalent cases. HIV+ MSM in ATHENA (black dots) and model prediction (thick line). Predicted number (thin line) and proportion (dashed line, right axis) of cases that are unaware of their infected status.

Results

Figure 2a and b shows the model curves that fitted best to the observed time series of annual new diagnoses and AIDS diagnoses (data on AIDS at diagnosis not shown). Figure 2d shows the estimated absolute number of new infections per year in the Netherlands. This peaked in 1983 with 802 new infections, and in 2004 with 554 new infections.

Estimates for the risk behaviour rate \( \beta(t) \), the reproduction number \( R(t) \) and the mean time to diagnosis are shown in Fig. 2c and d. Over the initial period (1980–1983), the estimate for the reproduction number \( R(t) \) is 2.39 [95% CI (confidence interval) 2.17–2.76]. Between 1984 and 1995, the risk behaviour rate declined by 2.3-fold (95% CI 2.03–2.83), indicating large reductions in risk behaviour, and thereby reduced the reproduction number \( R(t) \) below one to 0.89 (95% CI 0.85–0.93), that is, just below the epidemic threshold.
In a sensitivity analysis, the results on the key outcomes $\beta(t)$ and $R(t)$ appear to be very robust to a wide range of model variants. In particular, model results were consistent when assumptions about the relative infectiousness of disease stages, effect of diagnosis on behaviour, and time from diagnosis to start of therapy were varied. In all model variants, $R(t)$ for 2000–2004 is estimated to be near or above the critical threshold ($R = 1$), thus implying uncontrolled epidemic spread, with estimates of the current reproduction number ranging between 0.95 and 1.33, depending on the scenario (details available on request).

**Discussion**

The joint effect of HAART and risk behaviour on HIV incidence has been previously studied using mathematical models and empirical data [16–21,35]. Although based on different assumptions, all these studies come to the same conclusion regarding the potential for an increase in risk behaviour to offset the benefits of HAART in reducing transmission. Our study provides new evidence that this has actually occurred and quantifies its magnitude and timing within a well studied population of MSM.

A key feature of our study is the existence of several national databases recording diagnoses of HIV infection and AIDS, and deaths, allowing the diagnosis rate to be estimated reliably by simultaneously fitting to these time series within a robust inference framework. We were thus able to confirm that there has indeed been a recent increase in the diagnosis rate, reflecting a more frequent diagnosis rate, among MSM in the Netherlands, in spite of earlier work showing that the only way to reverse epidemic spread, and get to ‘serosort’ [40] are not proving effective at the population level and have not offset epidemic spread.

The introduction of HAART was accompanied by a decrease in the percentage of resistant strains among new infections [33,41]. However, the recent increase in annual new infections could in turn result in an increasing absolute number of resistant infections [42].

The widespread use of HAART has led to large reductions in AIDS morbidity and mortality (Figs. 2 and 3). Sustaining these reductions into the future will require either further improvements in treatment efficacy or a response to limit resurgent epidemic spread.

In conclusion, there is an increase in HIV transmission among MSM in the Netherlands, in spite of earlier diagnosis and subsequent effective treatment. The most effective intervention is to bring risk behaviour back to pre-HAART levels.

**Acknowledgements**

D.B. was supported by grant 7014 from AIDS Fund Netherlands and by a travel grant from NWO (Netherlands Organisation for Scientific Research). C.F. is funded by the Royal Society. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

**References**


