Evaluating the impact of antiretroviral therapy on HIV transmission

Joshua A. Salomon\textsuperscript{a,b,c} and Daniel R. Hogan\textsuperscript{b,c}

As global efforts proceed to scale up the delivery of antiretroviral therapy (ART) to HIV-infected persons in most urgent need, it is essential to understand the potential impact of treatment expansion on the transmission of new HIV infections. In this study, we use a series of simple mathematical models to explore the direction and magnitude of treatment effects on the sexual transmission of HIV. By defining the circumstances under which ART can reduce the number of new infections transmitted by treated patients, we provide critical benchmarks to aid in prioritizing efforts to maximize the population health impact of treatment and in evaluating the performance of different treatment programmes. We find that, based on the best currently available evidence of possible treatment effects on patient infectiousness, survival and behavior, the potential remains for either positive or negative changes in overall transmission. In relation to the total number of expected secondary infections caused by each infected person, however, these net treatment effects are relatively modest, particularly if treatment is initiated at advanced stages of the disease. This finding implies that treatment alone should not be expected to alter the population-level incidence of new infections dramatically, in the absence of changes in other factors including possible behavioral responses among uninfected persons and among infected persons who are not yet treatment candidates.

Introduction

Since the dramatic clinical benefits of HAART for HIV-infected patients were demonstrated more than a decade ago [1–3], the widespread application of HAART in various high and middle-income settings has been accompanied by substantial declines in AIDS mortality, as reported in north America, Europe, and Brazil [4–6]. Early observations that HAART could greatly improve patient outcomes in resource-poor settings [7] were subsequently confirmed in several small cohort studies [8–10]. Recent evidence from Zambia suggests that real-world HAART programmes are feasible and beneficial [11]. Whereas higher mortality in the first few months of treatment has been observed in developing countries compared with outcomes in north America and Europe, the subsequent response to HAART appears similar across low and high-income settings [12,13]. Treatment coverage in developing countries has advanced more slowly than in the industrialized world. Nevertheless, momentum for scaling up treatment in resource-poor settings has risen recently; the number of people receiving antiretroviral therapy (ART) globally increased from an estimated 400,000 at the end of 2003 to an estimated 2 million people by December 2006 [14–16]. Since 2006, the World Health Organization and partners in global HIV/AIDS control efforts have called for universal treatment coverage by the year 2010 [17]. Whereas treatment scale-up has the potential to reduce rates of HIV/AIDS-related mortality significantly, efforts to prevent new infections remain a critical priority. In 2007, an estimated 2.5 million people were newly infected with HIV [18]. A number of studies of the likely trends in the epidemic under various combinations of
prevention and treatment efforts have concluded that only a comprehensive response will yield sustained progress [19,20]. Given the urgent need to reduce the incidence of HIV infections, important questions emerge about the potential role of treatment in diminishing the sexual transmission of HIV. The possible effects of HAART on HIV transmission are complicated by a range of factors, acting in divergent ways. First, successful treatment has the direct effect of reducing a patient’s viral load and, consequently, the likelihood of transmitting the infection by sexual contact [21]. Second, effective treatment prolongs life and may therefore produce a net increase in the total number of sexual contacts for a treated patient, all else being equal. Treatment-associated changes in sexual behavior would further affect transmission dynamics, in either a positive or negative direction [22]. Finally, treatment availability has the potential for beneficial indirect effects at the population level by facilitating the uptake of prevention programmes such as HIV testing [23]; or detrimental indirect effects resulting from behavioral disinhibition, as suggested in a study of men who have sex with men (MSM) in San Francisco [24].

As no experimental study could be undertaken to investigate all possible dimensions of this question, and as decisions must inevitably proceed before all uncertainties are resolved, the potential net impact of scaled-up treatment on HIV incidence has been explored in a range of applications of mathematical models [19,20,25–37]. Previous studies have addressed specific issues such as the differential effects of treatment introduced at different epidemic stages [26], the conditions needed for long-run disease eradication [32], and the expected increase in drug resistance under high coverage rates [28]. Many of these earlier efforts have focused on predicting epidemiological trends in specific settings such as Australia [31], South Africa [33,36], Botswana [20], Uganda [30], India [20,35] and Thailand [37]. Although divergent predictions have emerged from these studies, the complexity of the mathematical models and the numerous assumptions and parameters contained therein defy straightforward comparison; deciphering the differences in results is thus challenging. Our primary aim in this paper is therefore to present simple heuristics to characterize the direction and likely magnitude of changes in HIV transmission in relation to treatment, under various scenarios. We aim to complement previous efforts by offering clear benchmarks that derive from straightforward mathematical computations, and by presenting a general framework that easily accommodates outcome analyses in various settings and under a wide variety of alternative assumptions.

We define a conceptual framework for evaluating the impact of ART on the sexual transmission of HIV infection, operationalize the approach using a series of simple mathematical models, and draw insights into potential population-level effects of treatment on the incidence of new HIV infections based on the best current empirical evidence on treatment outcomes.

### A simple model for characterizing net transmission effects in individuals

We begin with a simple framework for considering the potential influence of treatment on transmission at the individual level. This framework allows us to characterize the set of conditions under which treatment will have a neutral or positive effect on the number of new infections in terms of three critical dimensions: the reduction of a patient’s infectiousness by suppressing viral load through treatment; the number of years of life added by treatment; and changes in sexual behavior among treated patients, compared with a counterfactual of no treatment.

Taking first the simplified example of an infected person who has a series of single sexual contacts with multiple susceptible partners, the net change in secondary infections may be expressed simply as the product of three factors reflecting changes in infectiousness, survival and sexual behavior in relation to treatment status. The linear formulation results from a binomial model of infection, which is a mathematical representation of infection risks that follows from treating sexual contacts as independent, random ‘trials’, analogous to a series of (weighted) coin flips. If the probability of an infected person transmitting the infection to a susceptible partner in a single (unprotected) sexual contact is , then the number of expected secondary infections is simply , where is the number of susceptible partners. If we further assume that susceptible partners are acquired at a constant rate over some survivorship duration , then this expression can be recast as the product of three factors: . This formulation is commonly used in mathematical models of infectious disease dynamics and is referred to as the basic reproductive number, [38].

In this framework, we may represent the effects of treatment on each of the three components in this expression as a corresponding set of three scalar factors; that is, let us assume that treatment lowers the probability of transmission per act by a factor , reflecting reduced infectivity via the suppression of viral load; increases survivorship by a factor , and changes the rate of acquiring new susceptible partners by a factor , (reflecting for example the balance between reduced morbidity and increased counseling). The number of new infections in a treatment scenario is then , and the net impact of treatment on transmission will depend on whether the product is greater than or less than 1.
Using this insight we can construct a simple heuristic to describe the direction of change in the expected number of infections caused by an individual who begins treatment, based on any combination of values for the effects of treatment on infectiousness, survival and behavior. The impact of these three critical assumptions may then be summarized graphically in terms of a 'no-harm frontier' (Fig. 1). The no-harm frontier signifies the various combinations of values along these three dimensions under which the number of new infections transmitted by a particular treated patient will be less than or equal to the number of secondary infections in the absence of treatment. Mathematically, the no-harm frontier consists of all points in the three-dimensional space having an exact product of 1.

The logic of this heuristic is straightforward: if treatment doubles survivorship, then either transmission probabilities or the number of risky contacts must be halved to balance out the longer period of exposure. The other points on each of the lines result from analogous calculations.

Two main conclusions are highlighted by this simple arithmetic. First, advancing the survival benefits of treatment magnifies the need to maintain adherence at high enough levels to achieve durable suppression of the viral load, and also to ensure that the frequency of risky sexual behavior is diminished sufficiently to outweigh the rise in opportunities for risk-taking that accompanies longer survivorship. For example, if an individual on treatment survives for 8 years compared with one year without treatment and reduces risky sexual behavior by 50%, the probability of transmission per coital act would still need to be lowered by more than 75% to yield a net reduction in the expected number of future infections.

Second, it is possible to imagine combinations of the three factors that fall on either side of the 'no-harm frontier' based on the plausible values that each factor may take. The same information in Figure 1 may be recast slightly to present an alternative graphical display of the changes in behavior that are needed to offset any given combination of treatment effects on survivorship and transmission probabilities. In Figure 2, the variety of shaded areas demarcate ranges of behavior effects that would lead to a neutral treatment impact on secondary infections based on the levels of survival and changes in infectiousness on the horizontal and vertical axes, respectively. For example, any combination of survival and infectiousness assumptions residing in the lower right area in the figure would require at least a 75% reduction in risky sexual behavior to ensure that the number of new infections per treated patient does not increase; on the other hand, combinations of survival and infectiousness effects that reside in the upper left area would accommodate at least a doubling of risky sexual behavior without producing a net rise in new secondary infections.

The straightforward multiplicative example shown here may be extended to reflect other analogous questions, for example the probability that an infected person will transmit the infection to a long-term susceptible partner. These other applications require a modification of the basic model to allow for slightly more complicated computations based on a Bernoulli process model [39], but the mathematical results (not shown) are nearly identical to those presented here.

Fig. 1. ‘No-harm frontiers’ for treatment effects on transmission in relation to changes in infectiousness, survival and behavior. Each line represents a particular survival duration, from 2 to 12 years in 2-year increments.

Fig. 2. Levels of behavior change required to balance survivorship gains from treatment, accounting for changes in infectiousness under treatment. The shaded areas on the figure correspond to ranges of behavior change, expressed as percentage changes in the number of sexual contacts with susceptible partners compared with a no-treatment scenario.
Figures 1 and 2 offer simple tools for benchmarking different programmes in terms of the likely changes in transmission resulting directly from treatment scale up. These heuristics may be used to set clear and measurable targets for prospective monitoring of the performance of treatment programmes, or for evaluating existing programmes based on measured clinical and behavioral outcomes. In the next section, we review the current published evidence on the magnitude (and direction in the case of sexual behavior effects) of each of the three dimensions that delineate the ‘no-harm’ criteria, with an emphasis on findings from resource-limited settings where available.

**Evidence on treatment impact on infectiousness, survival and behavior**

**Infectiousness**

Studies have demonstrated a strong link between viral load and infectiousness [21, 40–42]. We therefore concentrate on treatment-induced changes in viral load in order to impute changes in probabilities of transmission. Recent studies of ART in resource-poor settings have demonstrated virological benefits comparable to those reported in the industrialized world (Table 1) [8–11, 43–62]. Observed reductions in viral load over 6–48 months of follow-up ranged from 1.5 to 3.6 log copies/ml while on HAART, and approximately 70% of patients reached an undetectable viral load (depending on the study population and how ‘detectable’ was defined). In Table 1, we translate the changes in viral load reported in each study into estimated changes in transmission probabilities based on the relationship reported in Quinn et al. [21]; computed reductions in per-contact transmission probabilities ranged from 74 to 96% across the 13 studies we reviewed containing information on viral load reductions (Table 1).

**Survival**

In terms of survival benefits, a recent study by Braitstein et al. [12] suggested that survivorship in the first few

<table>
<thead>
<tr>
<th>Table 1. Clinical and survival outcomes from selected studies of HAART in developing countries.*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Site</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Eleven countries [43]</td>
</tr>
<tr>
<td>Three continents [57]</td>
</tr>
<tr>
<td>Brazil [58]</td>
</tr>
<tr>
<td>Cambodia [44]</td>
</tr>
<tr>
<td>Cambodia [59]</td>
</tr>
<tr>
<td>Cameroon [45]</td>
</tr>
<tr>
<td>Cameroon [60]</td>
</tr>
<tr>
<td>Cote d’Ivoire [46]</td>
</tr>
<tr>
<td>Cote d’Ivoire [47]</td>
</tr>
<tr>
<td>Haiti [48]</td>
</tr>
<tr>
<td>Malawi [49]</td>
</tr>
<tr>
<td>Peru [50]</td>
</tr>
<tr>
<td>Senegal [9]</td>
</tr>
<tr>
<td>Senegal [50]</td>
</tr>
<tr>
<td>Senegal [51]</td>
</tr>
<tr>
<td>Senegal [52]</td>
</tr>
<tr>
<td>South Africa [8]</td>
</tr>
<tr>
<td>South Africa [61]</td>
</tr>
<tr>
<td>South Africa [62]</td>
</tr>
<tr>
<td>Thailand [53]</td>
</tr>
<tr>
<td>Thailand [54]</td>
</tr>
<tr>
<td>Uganda [10]</td>
</tr>
<tr>
<td>Uganda [55]</td>
</tr>
<tr>
<td>Zambia [11]</td>
</tr>
</tbody>
</table>

*Some figures have been estimated indirectly when exact figures were not provided in text or tables. Not all figures are based on the initial sample size (N) as a result of differing lengths of follow-up for study participants and limited use of CD4 cell count and viral load monitoring in some studies. For studies presenting different measures of survival, we have summarized intention-to-treat results that ignored loss to follow-up (details available from the authors).

<sup>1</sup>Calculated from median reduction in viral load column, based on Quinn et al. [21] estimate of 2.45-fold reduction in infectiousness for every 1 log_{10} decrease in viral load.

<sup>2</sup>For baseline CD4 cell counts (×10^6 cells/l) of ≤20 (first row), 21–50, 51–100, and >100 (range across these three groups in second row).

<sup>3</sup>For baseline CD4 cell counts (×10^6 cells/l) of ≤50 (first row) and >50 (second row).

<sup>4</sup>Results summarized here for adults and adolescents (>13 years old).

<sup>5</sup>Mean rather than median.

<sup>6</sup>Probability of remaining in care among those who returned for at least one follow-up visit.

<sup>7</sup>For baseline CD4 cell counts (×10^6 cells/l) of ≤50 (first row), 50 to <200, 200 to <350 and ≥350 (range across these three groups in second row).
months of therapy is lower in resource-poor settings than in developed countries, but that subsequent mortality is similar. In the papers we reviewed from developing countries, one-year survival ranged from 70 to 98% across studies, with much of the variation corresponding to differences in CD4 cell counts at baseline (Table 1). The only two studies that have followed patients for more than 1–3 years reported 75% [52] and 91% [56] survival after 5 years on HAART. Longer-term survival outcomes have been estimated via modeling exercises that leverage data from developed country cohorts, suggesting that median survivorship may range between 5 and 15 years depending on patient disease status at treatment initiation and intensity of care [4,63,64].

**Behavior**

The introduction of effective AIDS treatment in Europe and the United States has had mixed effects on sexual behavior. For example, the Swiss HIV Cohort Study (N = 4723) found no association between unsafe sex and optimal viral suppression for individuals on HAART [65]. Studies of MSM have, however, reported increases in unsafe sex with reductions in viral load [66] and increases in the prevalence of sexually transmitted infections among those receiving HAART [67]. Another concern is the possibility of increased risky sexual behaviour among seronegative individuals in response to the availability of ART. Sexual disinhibition has been most thoroughly documented via increasing rates of gonorrhoea and unprotected sex among MSM in San Francisco [24]. A meta-analysis by Crepaz et al. [22] found no significant association between the probability of having unprotected sex and either being on HAART or having an undetectable viral load. The study also found, however, that individuals who believed HAART protected against HIV transmission were more likely to practise unsafe sex.

At the time of publication, we are aware of only four studies presenting data on sexual behavior among those on HAART in resource-poor settings. Bunnell et al. [68] found that self-reported non-use or inconsistent use of condoms with seronegative or unknown-serostatus partners fell by 70% in the 6 months after HAART initiation in a rural Ugandan cohort who received counseling on sexual behavior along with treatment. In a cross-sectional study in Kampala, Uganda, Bateganya et al. [69] observed that recipients of HAART were not significantly more likely to report being sexually active than non-recipients [adjusted odds ratio (OR) 2.0 (0.3–9.9)]; at the same time, HAART recipients were more likely to report consistent condom use, disclose their HIV serostatus to their partners and report treatment for a sexually transmitted disease in the past 6 months. In Côte d’Ivoire, Moatti et al. [70] reported lower levels of self-reported unprotected sexual intercourse in the previous 6 months among those receiving HAART [adjusted OR 0.52 (0.29–0.93)]. Finally, Chen et al. [71] found no change in the relative odds of reporting inconsistent condom use for individuals with HAART experience compared with those without.

On the basis of the information reviewed here, we may revisit Fig. 2 to consider what conclusions may be drawn about behavior changes needed to balance the currently observed changes in infectiousness and survival as a result of treatment. Limitations in the available evidence base imply fairly wide ranges in both of these dimensions. For example, we may imagine a rectangular area in Figure 2 bounded by the range of infectiousness changes inferred from the studies in Table 1 (reductions between 74 and 96%) on the vertical, and by a more uncertain range of survival derived from previously described modeling studies (5–15 years) on the horizontal. Given these ranges, we note a wide range of uncertainty in the level of required behavior change. A combination of relatively modest survival gains with relatively high effectiveness in suppressing transmissibility would offset a twofold increase in risky sexual behavior; on the other hand, more optimistic survival outcomes combined with more limited reductions in transmissibility might require up to a fourfold decrease in risky behavior to avoid any rise in secondary infections.

**Estimating the magnitude of treatment effects on secondary transmission**

Finally, we extend our analysis to consider the impact of treatment on transmission at the population level by evaluating the fraction of sexual transmission attributable to the pool of treatment candidates in illustrative case examples reflecting different risk groups and treatment eligibility scenarios.

We defined an array of cases to represent different types of individuals newly infected with HIV, characterized by particular patterns of sexual behavior. Results are shown for the two most extreme cases. The first represents an individual who has only one susceptible partner for the remainder of his or her life (e.g. a married person with an initially uninfected spouse). The second represents an individual who has 500 new susceptible partners per year, with each partnership consisting of one coital act (e.g. a sex worker). For simplicity, we modeled only serodiscordant partnerships.

In order to evaluate the likely impact of treatment within the broader context of the total number of infections expected over the remaining lifetime of an infected person, we modeled three different treatment scenarios: no treatment; treatment initiated at CD4 cell counts of $200 \times 10^6$ cells/l or less, in line with World Health Organization recommendations for resource-poor settings [72], and treatment initiated at CD4 cell counts of...
50 \times 10^6 \text{ cells/l} or less. To allow for transmission probabilities that may vary over the course of infection, we partitioned the natural history of infection into stages representing early infection, the long intermediate stage of asymptomatic HIV (further stratified by months since infection), and AIDS (further stratified by time before death), and modeled the expected number of secondary infections during each stage. For simplicity in presentation, we refer to ‘early’ and ‘late’ AIDS stages, mapping approximately to CD4 cell counts between 200 and 50 \times 10^6 \text{ cells/l}, and CD4 cell counts less than 50 \times 10^6 \text{ cells/l}, respectively. The expected numbers of new infections in each stage were computed based on stage-specific levels of infectiousness, durations and associated sexual behavior (Table 2) [21,64,73,74]. Calculations are described in Box 1. Our analysis of infections by stage extends the approach applied in a previous analysis by Hayes and White [75], who modeled expected infections over three stages, without reference to treatment, in order to draw conclusions about the importance of primary infection for transmission.

Transmission probabilities by stage were derived from a recent study by Wawer et al. [73]. To accommodate the reported stratification of late-stage infection by time to death in the study by Wawer et al. [73], we assumed that ‘early’ AIDS in our model corresponds to the period between 6 and 25 months before death, whereas ‘late’ AIDS corresponds to the last 6 months before death. For the effect of treatment on infectiousness, we assumed that HAART would reduce the viral load by 2.8 orders of magnitude (the median from Table 1). Quinn et al. [21] estimated a 2.45-fold change in transmission rates for every log change in viral load. We applied this rate ratio to per-act transmission probabilities (as the error in this approximation is negligible for probabilities of this order of magnitude), which resulted in an estimated 92% reduction in infectiousness (computed as 1 – 1/2.45^{2.8}). Taking early-stage AIDS as an example, this assumption implies a per-act transmission probability under treatment of 0.00026 (computed as 0.0032 \times 0.08).

Although long-term experience with HAART in resource-poor settings is limited, we assumed that initiating treatment when CD4 cell counts were between 200 and 50 \times 10^6 \text{ cells/l} would yield 8.5 years of survivorship and when CD4 cell counts were below 50 \times 10^6 \text{ cells/l}, 5.5 years of survivorship, following

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline value</th>
<th>Range*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission probability per act [73]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early HIV infection, (0–4 months since infection)</td>
<td>0.0082</td>
<td>0.0039–0.0150</td>
</tr>
<tr>
<td>Asymptomatic HIV, by months since infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–15</td>
<td>0.0015</td>
<td>0.0002–0.0055</td>
</tr>
<tr>
<td>16–35</td>
<td>0.0010</td>
<td>0.0000–0.0054</td>
</tr>
<tr>
<td>36–83</td>
<td>0.0007</td>
<td>0.0000–0.0011</td>
</tr>
<tr>
<td>84–93</td>
<td>0.0014</td>
<td>0.0002–0.0050</td>
</tr>
<tr>
<td>Untreated AIDS, by months before death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–25</td>
<td>0.0032</td>
<td>0.0015–0.0061</td>
</tr>
<tr>
<td>6–15</td>
<td>0.0043</td>
<td>0.0018–0.0084</td>
</tr>
<tr>
<td>Reduction in coital frequency, relative to pre-AIDS stages (%)d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated AIDS, by months before death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–25</td>
<td>13</td>
<td>7–26</td>
</tr>
<tr>
<td>6–15</td>
<td>38</td>
<td>19–76</td>
</tr>
<tr>
<td>Treated AIDS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td>13</td>
<td>0–50</td>
</tr>
<tr>
<td>Late</td>
<td>38</td>
<td>0–30</td>
</tr>
<tr>
<td>Relative reduction in transmission probabilities on treatment (%) [21]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival on treatment (years) [64]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early treatment</td>
<td>8.5</td>
<td>4.3–17</td>
</tr>
<tr>
<td>Late treatment</td>
<td>5.5</td>
<td>2.8–11</td>
</tr>
</tbody>
</table>

*Ranges used in one-way sensitivity analyses, except for ranges around transmission probabilities, used jointly in multivariate, Monte Carlo uncertainty analysis. Ranges for transmission probabilities based on 95% confidence intervals in Wawer et al. [73]. Ranges for reductions in coital frequency in untreated AIDS, at 16–25 or 6–15 months before death reflect halving and doubling of baseline values. Ranges for reductions in coital frequency for treatment were arbitrarily chosen to include a broad range of possible treatment effects on behavior, reflecting lack of evidence to date. Ranges around relative reductions in transmission probabilities on treatment calculated on the basis of high and low values for changes in viral load observed across studies in Table 1.

Although the period of early infection is labelled as months 0–5 after seroconversion in the tables presented by Wawer et al. [73], the duration of the period is explicitly noted as 5 months elsewhere in the paper, which implies a period spanning months 0–4. We adopt the 5-month duration and append the additional month to the start of the second period for purposes of transmission calculations. The period from months 16 to 83 maps to transmission probabilities estimated from the group labelled ‘prevailing index partners’ in Wawer et al. [73], with the duration imputed by subtracting durations for other specified stages from an assumed total duration of 10 years from infection to death in the absence of treatment. The period from months 84–93 maps to the period 26–35 months before death in Wawer et al. [73].

1In the baseline, we assume that there is no transmission in months 0–5 before death. In sensitivity analyses, we consider alternative assumptions whereby the transmission probability in this period is the same as in the first 5 months of infection, and coital frequency is the same as in months 6–15 before death.

2Coital acts per month in pre-AIDS stages assumed to be 10 for the single-partner case [73] and 25 for the multiple-partner case [74].
previous studies [58]. Durations of other disease stages were defined to match the periods over which Wawer et al. [73] estimated infectiousness (Table 2).

We conducted one-way sensitivity analyses on key parameter values to examine the robustness of our conclusions to particular assumptions. For a simple case in which transmission probabilities and sexual behavior are constant throughout the lifetime of an infected person, the lifetime probability of transmission to a susceptible partner would be expressed as follows:

$$\Pr[\text{Infection}] = 1 - (1 - p)^k$$

where $p$ equals per-coital-act transmission probability; $k$ equals the duration of partnership (in months); $a$ equals contacts per month.

This basic expression may be expanded to accommodate variation in transmission probabilities and behavior across stages, and to partition the expected number of infections by stage, as follows:

$$\Pr[\text{Infection}_t] = \left[1 - (1 - p_t)^{k_t a_t}\right] \times \prod_{j=0}^{t-1} (1 - p_j)^{k_j a_j}$$

where $p_t$ equals the per-coital-act transmission probability during stage $t$; $k_t$ equals the duration of stage $t$ (in months); $a_t$ equals contacts per month during stage $t$.

Figure 3 shows, for both extreme behavioral case examples, the expected number of infections produced by the HIV-infected case by stage and under different treatment scenarios, expressed relative to the total numbers of expected infections in the no-treatment scenario. The balance of infections across stages differs in the two behavioral cases, because in the single-partner case, the probability of the partner remaining susceptible declines over time as a result of cumulative risk exposure, which concentrates infections in the early stages; in the multiple-partner case, by contrast, new susceptible partners are acquired continuously, which spreads secondary infections over the progression through more advanced stages. In the absence of treatment, approximately 14% of total transmission occurs once an individual reaches the treatment-eligible early-AIDS stage in the single-partner case, compared with approximately double this fraction in the multiple-partner case.

![Fig. 3. Expected numbers of lifetime infections in different treatment scenarios, by stage of infection.](image)

(a) The case example of an infected person with a single susceptible, lifelong partner. (b) The case example of an infected person with 300 susceptible, single-contact partnerships per year. In each figure, the top bar displays the ‘no treatment’ scenario, the middle bar represents treatment initiated at CD4 cell count levels between 200 and 50 × 10^6 cells/l, the bottom bar represents treatment initiated at CD4 levels of 50 × 10^6 cells/l or below. □ Early HIV; □ asymptomatic HIV; ■ untreated AIDS; □ treated AIDS.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
Based on the timing of treatment initiation, either positive or negative treatment effects on transmission are possible. In the early-treatment scenario (assuming treatment begins at CD4 cell counts between 200 and $50 \times 10^6$ cells/l), the expected number of secondary infections declines by approximately 7% compared with no treatment in the single-partner case, and by 17% in the multiple-partner case. On the other hand, if treatment is initiated later (at CD4 cell counts below $50 \times 10^6$ cells/l), the expected number of secondary infections rises by 3% in the single-partner, and by 8% in the multiple-partner case.

Results of sensitivity and uncertainty analyses indicate that our qualitative findings are relatively robust to variation in the parameter values used. The direction and general magnitude of treatment effects in our baseline analyses persist when we vary the duration of survivorship on treatment, risky behavior in late-stage compared with early-stage disease, or the behavioral effects of treatment, across the full ranges indicated in Table 2. In sensitivity analyses on these variables, effects range from reductions of 2–11% or 5–24% as a result of early treatment in the single-partner or multiple-partner cases, respectively; or increases of 1–3% or 4–16% as a result of late treatment in the single or multiple-partner cases, respectively. Doubling or halving all transmission probabilities simultaneously produces no change in the relative effects of treatment in the multiple-partner case because infection numbers are linear in these rates. In the single-partner case, doubling or halving all transmission probabilities results in net treatment effects ranging from 2–12% reductions as a result of early treatment or from 1–5% increases as a result of late treatment. Allowing each stage-specific transmission probability to vary independently in a multivariate, Monte Carlo uncertainty analysis likewise yields only modest deviations from baseline estimates of relative transmission effects; 95% uncertainty intervals range from 2–9% or 9–23% reductions as a result of early treatment in single or multiple-partner cases, respectively; and from 1–3% or 3–11% increases as a result of late treatment.

There are two instances in which varying assumptions can lead to differences in the direction of estimated treatment effects on transmission. First, if we consider the lower bound on the estimated infectiousness effect (i.e. assume a 74% reduction in per-act transmission probabilities, compared with 92% in the baseline), even treating individuals early could result in a higher expected number of lifetime infections compared with no treatment. At this extreme, infections increase by 3% compared with no treatment in the single-partner scenario, and by 11% in the multiple-partner scenario. Second, if we assume in the no-treatment scenario that sexual transmission occurs in the final 6 months of life, contrary to the base-case assumption that no transmission occurs at this late stage, then even late-stage treatment can reduce the total number of lifetime infections. For example, if behavior in the final 6 months of life were the same as in the preceding 10 months, and infectiousness in this stage matched that in early infection, then initiating treatment for late-stage AIDS would produce a net transmission decrease of 2% in the single-partner, and 8% in the multiple-partner case.

Discussion

Although uncertainties persist about the potential impact of ART on HIV transmission, efforts to reduce the burden of HIV/AIDS must proceed and should be informed by the best available information at this time. Mathematical models can provide useful inputs into these policy decisions by synthesizing available information to predict consequences that are not directly observable. In this study, we used a series of simple mathematical models to devise heuristics regarding the potential direction and magnitude of effects of treatment on transmission, defining the circumstances under which ART can reduce the number of new infections transmitted by treated patients. We thus aim to provide critical benchmarks that help prioritize efforts to maximize the population health impact of treatment as coverage expands, and help evaluate the performance of different treatment programmes in terms of their expected effects on transmission.

Based on the best currently available evidence of possible treatment effects on patient infectiousness, survival and behavior, we find that the potential remains for either positive or negative changes in overall transmission. Transmission effects will be more favorable if treatment is initiated at an earlier stage of the disease. It is worth noting, however, that in relation to the total number of expected secondary infections caused by each infected person, the net effects of treatment, either positive or negative, are likely to be relatively modest, particularly when treatment is initiated at advanced stages. For example, if treatment is initiated at CD4 cell counts of approximately $200 \times 10^6$ cells/l, we estimate that expected transmission would be reduced by approximately 7% for a person with a single susceptible, life-long partner or by 17% for a person with a large number of single-contact partners. If, on the other hand, treatment is initiated at CD4 cell counts of approximately $50 \times 10^6$ cells/l, we expect transmission to increase by 3% in the single-partner case or by 8% in the multiple-partner case.

Other investigators have applied dynamic transmission models to examine the potential population-level impact of antiretroviral scale-up [20,25–33,35]. Results have spanned a wide range of estimated effects. For example, Abbas et al. [26] estimated that treatment introduced early in an epidemic, administered to 100% of AIDS cases, could reduce cumulative new infections by 29% over 10 years. Gray et al. [30], using a simulation model
parameterized based on empirical data from studies in Rakai, Uganda, predicted that reductions in incidence as a result of treatment could range from 9 to 19%, with the difference relating to the magnitude of treatment-related reductions in viral load observed in two different Rakai studies. Garnett et al. [29] simulated the impact of ART under two alternative cases in which individuals either begin treatment at a relatively late stage of disease or are recruited actively at earlier stages. In that analysis, predicted declines in incidence over a 10-year period after treatment introduction were approximately 20% in the ‘passive’ case or more than 50% in the ‘active’ case.

The simple models presented here are not intended as a substitute for these more sophisticated models, which aim to describe the long-term trajectory of HIV epidemics under specific control strategies, or with reference to particular settings. Nor do we intend to inform detailed clinical practice questions, for example regarding the timing of treatment and use of diagnostic tests [76]. Rather, we offer general guidance on how the net effect of treatment on transmission will move systematically in relation to three specific aspects of treatment outcomes: reduced viral load; increased survivorship; and changes in patient risk behavior. The models presented here produce rules of thumb that are operationalized in stylized case examples to map from individual patient outcomes into likelihood population-level effects. As such, our study complements previous efforts by offering a simple mathematical framework to identify general conclusions applicable in various settings, and also to accommodate outcome analysis under a wide variety of alternative assumptions and scenarios. In this paper we identify plausible ranges for key model parameters based on evidence available at this time, but inevitably are limited by data that are imperfect and incomplete. As experience with antiretroviral programmes in developing countries accumulates, we expect that the specific numerical results in this paper may be revisited and refined.

Nevertheless, certain key lessons emerge. The finding that either positive or negative changes in secondary transmission remain plausible under the currently observed effects of treatment on infectiousness, survival and behavior reinforces the need to ensure that treatment programmes sufficiently emphasize reducing risky sexual behavior in patients at the same time as they aim to optimize clinical outcomes. Nevertheless, the simple examples here suggest that current criteria for initiating treatment in resource-poor settings will probably minimize the overall impact of treatment on transmission, as only a fraction of transmission arises from the treatment-eligible population based on prevailing guidelines. The relatively modest size of the net changes in transmission attributable to treatment, in relation to the total numbers of secondary infections per person, implies that treatment alone should not be expected to alter the population-level incidence of new infections dramatically, without changes in other factors. Efforts to promote positive behavior change among uninfected persons and among infected persons who are not yet treatment candidates, therefore remain critical priorities in the global response to HIV/AIDS epidemics.

Sponsorship: The authors gratefully acknowledge funding support from the National Institute on Aging and the Bill and Melinda Gates Foundation.

Conflicts of interest: None.

References


17. UN General Assembly. 60/262. Political declaration on HIV/AIDS. Washington, DC, USA: UN; 2006.


