PLOS | MEDICINE

Investigating the Impact of Treatment on New HIV Infections

www.ploscollections.org/TasP2012

Produced with support from the HIV Modelling Consortium, which was funded by a grant from the Bill & Melinda Gates Foundation.

Image Credit: Psyberartist at flickr.com, Manta Ray Media/HIV Modelling

PUBLIC LIBRARY of SCIENCE www.plos.org Volume 9 | Issue 7 | July 2012

PLoS MEDICINE

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Models, Data, and Questions—Towards Evidence-Based Decision-Making

The HIV Modelling Consortium Treatment as Prevention Editorial Writing Group*

Abstract: Antiretroviral therapy (ART) for those infected with HIV can prevent onward transmission of infection, but biological efficacy alone is not enough to guide policy decisions about the role of ART in reducing HIV incidence. Epidemiology, economics, demography, statistics, biology, and mathematical modelling will be central in framing key decisions in the optimal use of ART. *PLoS Medicine,* with the HIV Modelling Consortium, has commissioned a set of articles that examine different aspects of HIV treatment as prevention with a forward-looking research agenda. Interlocking themes across these articles are discussed in this introduction. We hope that this article, and others in the collection, will provide a foundation upon which greater collaborations between disciplines will be formed, and will afford deeper insights into the key factors involved, to help strengthen the support for evidence-based decision-making in HIV prevention.

Introduction

The 19th International AIDS Conference will meet in Washington, District of Columbia, 22–27 July 2012. Since the last International AIDS Conference in Vienna two years ago, more than five million people globally have become newly infected with HIV [1,2]. In South Africa, a country with one of the largest HIV epidemics, 3% of the young men and women who were 19 years old and uninfected at the time of the last conference will now be infected [3]. Indications that the rate of new HIV infections in several countries may have declined recently are extremely welcome. Moreover, the recent UNAIDS Investment Framework [4] and President's Emergency Plan for AIDS Relief guidance on combination prevention [5] suggest that combining existing interventions and scaling them up could have further significant impact on reducing HIV incidence. However, these strategies are not expected to bring the epidemic fully under control.

Advances in HIV prevention research over the past two years have generated considerable optimism. First, it was shown that a 1% tenofovir vaginal microbicide gel reduced HIV acquisition in women in South Africa [6], and this was followed by a trial demonstrating that daily oral co-formulated tenofovir and emtricitabine reduced the risk of HIV acquisition in men who have sex with men (MSM) [7]. Subsequently, daily oral tenofovir alone or combined with emtricitabine was shown to reduce the risk of HIV acquisition in heterosexual men and women in long-term relationships in Uganda and Kenya [8]. There have also been some indications that a vaccine candidate (RV144) provides some short-term protection against infection [9]. These modalities provide a partial reduction in risk, but some studies on pre-exposure prophylaxis have produced conflicting results, highlighting that many questions in this field remain unanswered [10].

However, the finding that has created that greatest excitement has been that HIV-infected individuals who are given antiretroviral therapy (ART) are much less likely to transmit the infection to their heterosexual partners than those who are not. This finding was shown in the HPTN 052 trial [11] (Box 1), which was chosen as the *Science* magazine breakthrough of the year for 2011 [12]. If viral load is fully suppressed, those on ART may effectively be almost uninfectious. Although anticipated [13,14], this finding has catalyzed enormous interest in how ART could not only benefit the individual provided with the medicines, but also reduce the epidemic burden of the communities in which they live by limiting HIV transmission.

The role of ART in reducing HIV incidence will probably be among the most important topics in the field of HIV prevention for years to come, and it is already being debated urgently at national and international levels, within major normative agencies and charities, and by donors and implementers. The issues cut across the domains of epidemiology, economics, statistics, demography, virology and immunology, behavioural science, mathematical modelling, and clinical trials, and demand an interdisciplinary approach.

The HIV Modelling Consortium aims to coordinate and promote research across these disciplines and streamline commu-

Citation: The HIV Modelling Consortium Treatment as Prevention Editorial Writing Group (2012) HIV Treatment as Prevention: Models, Data, and Questions—Towards Evidence-Based Decision-Making. PLoS Med 9(7): e1001259. doi:10.1371/journal.pmed.1001259

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: The HIV Modelling Consortium is funded by the Bill & Melinda Gates Foundation through a grant to Imperial College London. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: GG is currently employed by the Bill & Melinda Gates Foundation. In the past 5 yeas GG has received grants from the Wellcome Trust, Medical Research Council the Gates Foundation and UNAIDS. In the past 5 years GG has been a consultant for GSK and Merck working on HPV vaccination and for Viropharma to work on stem cell transplantation. CM is employed by a commercial company that provide consulting services in health economics to pharmaceutical companies and donor organizations. Much of our work is focussed on HIV/AIDS and we have been involved in projects that evaluated the health economic impact of ARV treatment for prevention. The company was, however, not compensated for being part of the HIV modeling group that wrote this article. WM receives US NIH grant support. JS is a member of the *PLoS Medicine* Editorial Board. BZ is PI of the ALPHA network of HIV community-based cohort studies that may supply data to the modelling consortium. All other authors have declared that no competing interests exist. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the agencies represented, including US Centers for Disease Control and Prevention (CDC), PEPFAR, UNAIDS, WHO, or the Bill & Melinda Gates Foundation.

Abbreviations: ART, antiretroviral therapy; MSM, men who have sex with men

* E-mail: timothy.hallett@imperial.ac.uk

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Box 1. The HPTN 052 Trial

The HPTN 052 trial enrolled 1,763 HIV-1 serodiscordant couples (i.e., couples in which one partner is HIV-infected but the other is not) in which the CD4 cell count for the HIV-infected partner was between 350 and 550 cells/μl. The HIV-infected partners were randomized either to receive ART immediately ("early ART" arm) or to receive ART when their CD4 cell count dropped below 250 cells/μl ("delayed ART" arm). The couples were followed up for a median of 1.7 years, and substantial effort was made to ensure that viral suppression was achieved among those in the early ART arm. A total of 39 transmission events were observed. Genetic linkage analysis confirmed that 28 of these were linked to the stable partner. Of these 28 linked transmissions, 27 were in the delayed ART arm and one was in the early ART arm, resulting in an estimated 96% reduction (95% confidence interval: 73%–99%) in the risk of transmission from HIV-infected individuals on early ART compared with delayed ART. Earlier ART was also associated with significant improvement in a composite indicator of morbidity and mortality (41% [95% confidence interval: 12%–60%] reduction).

Although the HPTN 052 study was the first randomized controlled study to demonstrate the impact of ART on transmission, an earlier observational study among couples recruited for another trial had previously indicated that ART was associated with a 92% reduction in the risk of transmission [13]. Other observational studies also support that the risk of transmission when virally suppressed on ART is very substantially reduced [14,57]. However, many questions remain about the impact of ART on transmission, including the durability of the effect, levels of suppression that would be possible in other settings, and the impact through other routes of HIV transmission (especially unprotected anal sex).

nication between decision-makers and researchers. Mathematical modellers have considered the potential impact of ART on HIV incidence in a variety of scenarios and settings over the past 15 years, with model estimates becoming more refined as improved data have become available. A collaborative meeting of the HIV Modelling Consortium was held in November 2011 (http://www.hivmodelling.org/events/potential-impact-treatment-hiv-incidence) to review findings and identify priorities for future research. Several interlocking themes arose at the meeting, which are covered by the set of articles in this special collection, "Investigating the Impact of Treatment on New HIV Infections" (http://www.ploscollections.org/TasP2012) [15–23]. In this article we seek to set each piece in context, and describe important issues that are beyond the scope of this collection.

The Potential Impact of ART on HIV Incidence

Fundamentally, the impact that a treatment programme can have on preventing infections in an epidemic is determined by two main factors. First, it is determined by the number of onward transmissions generated by a newly infected person *before* they start treatment, which is determined by the biology of HIV infection, patterns of sexual contact between partners, the effects of other prevention interventions, and the rates of HIV testing and linking to care (Figure 1). Second, the impact is determined by the number of onward infections generated by an individual *after* ART initiation, which additionally depends on the biological efficacy of treatment, as well as adherence and retention in care. Estimating the population-level impact of expanded access to ART therefore involves synthesising diverse sources of information and managing substantial amounts of uncertainty about virology, immunology, human sexual behaviour, and the long-term performance of prevention programmes. The biological efficacy data provided by the HPTN 052 trial [11] is only one piece of this puzzle.

Mathematical models provide a framework within which to assemble this information, and several models of the epicentre of the worldwide epidemic, sub-Saharan Africa, have been developed and used to investigate the potential impact of treatment on HIV incidence. As different studies have addressed different questions and made different assumptions, it has been unclear whether or not these models fundamentally agree about the potential impact of particular treatment interventions in reducing HIV incidence. If they do, this could increase confidence in their collective findings, but if they do not, then this provides an important note of caution when considering results and highlights areas for further investigation.

A Systematic Comparison of 12 Models

In this collection, Eaton et al. [15] present the results of a systematic model comparison exercise in which 12 of these models were used to simulate the same sets of interventions. The model results were relatively consistent for short-term (eight-year) projections of reductions in incidence associated with treatment. For instance, if, hypothetically, 80% of individuals were treated after their CD4 cell count reaches 350 cells/μl (approximating current international guidelines; Box 2), the models projected that the incidence rate would be reduced by 35%–54% after eight years, compared with what the incidence would be in the absence of any ART. All models suggested that the existing treatment scale-up in South Africa should have already reduced new infections (incidence in 2011 is estimated to be 17%–32% lower than if there had been no ART [15]). The consensus that treatment provided within current guidelines has a prevention benefit is significant and should serve to reinforce the case for continuing to improve access to ART. However, there was much more variation in long-term (38-year) projections of reductions in incidence. One important way in which the models differ is in how they represent the behaviours leading to transmission, such as heterogeneity in sexual risk behaviours and patterns of contact with respect to age, which are notoriously hard to quantify [24]. Another difference is in how they represent the biology of infection, in particular the rate of CD4 cell count decline and relative infectiousness [25,26], about which there is little comprehensive agreement. It will be important to consider the influence of these factors on the key outcomes of interest when interpreting future modelling studies on this topic.

Connecting Model Projections to the "Real World"

When using extremely ambitious assumptions about the ability of ART programmes to test and start treatment of HIV-infected individuals very soon after infection, and retain them in care, five of nine models compared by Eaton et al. [15] suggested that incidence would be reduced by more than 90%, similar to the modelling predictions reported by Granich et al. [27]. However, these assumptions can be contrasted with recent real world experience in which the HIV testing rate was 52% in the cross-sectional, nationally representative South African National HIV Prevalence, Incidence, Behaviour and Communication Survey [3], and the repeat testing rate of individuals in an intensive community-mobilising intervention was 28% [28]. In addition, linkage from testing to ART uptake is assumed to be 100% in the

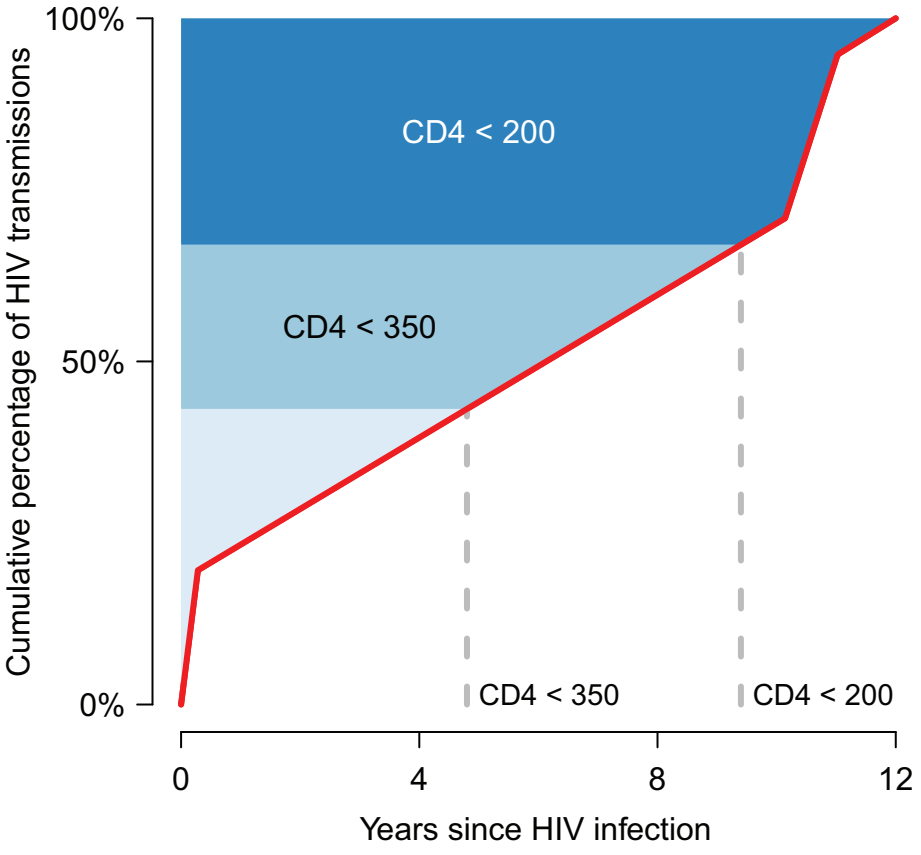

Figure 1. A framework for understanding the epidemiological impact of HIV treatment. The published results of models [38,53–55] that have estimated the contribution of different stages of HIV infection to onward transmission are compiled in a median cumulative distribution of infections generated by one infected person over the course of his/her infection in the absence of treatment (red line). The horizontal axis shows time from the time of infection to 12 years, which is the mean survival time for those with untreated HIV infection [56]. The vertical axis shows the cumulative transmission, from 0% (no new infections generated yet) to 100% (all onward transmission completed). (Note that the uncertainty in this distribution is not indicated.) The shading indicates the approximate CD4 cell count category at each time point [25,26]. Currently, treatment tends to be initiated well below a CD4 cell count of 200 cells/μl [32], meaning that the contribution of treatment to prevention is minimal because most of the transmission from that person has already occurred before treatment starts. If increased testing and improved linkages to care enabled individuals to start treatment at a CD4 cell count very close to 200 cells/μl, this could result in a substantial reduction in HIV incidence, because ~25%–30% of transmission normally arises from individuals after that point. The prevention impact would be expected to be even greater with initiation close to a CD4 cell count of 350 cells/μl. If the average number of new infections arising from an infected person in a susceptible population exceeds one before treatment could be feasibility initiated, then treatment could not eliminate the HIV epidemic. In this framework, the influence of other forms of prevention will be to change the shape of the graph. For instance, if many men are circumcised or individuals have fewer new sexual partners per time unit, then new infections arising from an infected person will grow more slowly over time, so that on average one new infection might be generated only after the point at which a feasible programme could have initiated treatment.
doi:10.1371/journal.pmed.1001259.g001

models, but has been about 33% in actual programmes [29]. Rather than 0% refusal of uptake of treatment, as assumed in the models, some settings have seen 20% refusal [30]. Finally, the dropout rate from programmes was 1.7% per year in the most optimistic model simulations presented in Eaton et al., compared with around 10% over the first year in the IeDEA network of clinics [31–33].

These inconsistencies between modelling assumptions and projects and real world situations do not mean that treatment cannot be used to generate greater reductions in incidence, but rather that major advances in programme coverage and delivery will be required to fully exploit the potential prevention benefits of treatment. These are operational barriers that could be improved without the development of new scientific prevention technologies, but which will nevertheless require substantial investment in health services.

In many models, including several of those in the modelling comparison [15], several significant simplifying assumptions about other factors that might influence success were made, because the exercise was focussed on the impact of a simple and stylized treatment programme on HIV incidence. In particular, most models did not explicitly include the relationship between adherence to ART regimens and degree of viral suppression, which would affect the therapeutic benefit, the prevention effect, and the potential for emergence of drug-resistant virus. Drug resistance is an important issue, especially over the long timescales considered here, because it effectively weakens the impact of existing first-line regimens and could cause greater reliance on second- and third-line treatment regimens, which are currently more expensive. There are many other considerations that the modelling comparison by Eaton et al. did not address, such as the interaction of ART with behavioural interventions and the best

Box 2. Current International Guidelines for Use of ART

The current World Health Organization international guidance recommends that HIV-infected patients with CD4 cell count ≤350 cells/μl be initiated on ART. In addition, patients with advanced clinical disease or HIV-infected people with active tuberculosis should be immediately initiated on ART, irrespective of CD4 cell count [58]. In April 2012, new guidance was issued that HIV-infected individuals with a long-term partner who is HIV-uninfected could also be considered for ART initiation [59]. New guidelines will be promulgated by the World Health Organization in 2013 [60].

However, these guidelines do not necessarily reflect the care that patients actually receive. National guidelines may or may not fully reflect the World Health Organization guidance, and typically, constraints on resources, the capacity of health systems, and care-seeking behavior result in individuals being initiated on ART at lower CD4 cell counts than the guidelines recommend.

use of diagnostic tools that could measure viral load or CD4 cell count at point of care, which are also the subject of ongoing research [34–37] but beyond the scope of this collection.

Evidence of Impact from Existing Programmes

Consensus across multiple models can be reassuring, but it is still possible that all the models could be wrong if, for instance, the small number of key data sources they rely on are not representative, or if all the models do not incorporate some crucial aspects of the system. Another essential check for models is a comparison of their projections with real data: in this case, the observed impact of treatment programmes in industrialised countries that have already achieved good access to treatment [38]. In this collection, Smith et al. [16] review the data that have been interpreted as showing that treatment has already had an impact on reducing incidence, showing apparent consistency between modelled expectations and reality. However, Smith et al.[16] advise caution when interpreting the level of evidence implied, particularly where indirect metrics for ART exposure (such as community viral load) and proxies for HIV incidence (such as new diagnoses) are used.

In this collection, Wilson [17] describes the examples of Australia and France, among other settings, where, despite high testing rates and coverage of treatment among MSM, HIV incidence has not decreased. This is in contrast to what models suggest should have occurred if the assumptions about treatment as prevention from heterosexual studies are applied to MSM populations. It will be essential for modellers to learn from the past by reconciling these and other observations to refine future model projections.

The Role of Early HIV Infection

One particular issue that may prevent even the most ambitious treatment programmes from reducing HIV epidemics to very low levels is the role of early HIV infection in sustaining HIV transmission. Early HIV infection covers the time shortly after infection—and usually before HIV diagnosis—when viral concentration in the blood spikes and individuals are more infectious [39]. If a substantial proportion of transmission occurs during early infection, the impact of treatment programmes will be less [40] (Figure 1). This could in part explain the apparent lack of preventive efficacy of ART in epidemics among MSM, as explored in Australia and elsewhere by Wilson [17] and in other examples examined by Kumi Smith et al. [16]. However, it is an open question whether early HIV infection is a dominant factor in sustaining epidemics in sub-Saharan Africa, and it has been argued that the contribution of acute infection to sustaining epidemics is not a primary determinant of the impact of treatment interventions. In this collection, Cohen et al. [18] debate the size and significance of this effect, and call for new data to be collected that may help this to be determined.

Economic Considerations

Ideally, public health policy should be driven by maximising improvements in the health of populations, rather than by economic considerations. But the HPTN 052 [11] findings have come at a difficult time for the public health response to HIV. After years of rapid growth, funding commitments and disbursements have stabilised or been reduced [41], and only a few countries in sub-Saharan Africa are currently able to achieve the high levels of treatment coverage for those eligible recommended by current international guidelines (Box 2) [1,2]. While the cost of providing treatment has fallen dramatically in recent years [42], offering ART to individuals who are not in immediate clinical need may continue to be significantly more expensive and complex than other existing methods for reducing HIV transmission, such as male circumcision [43] and some forms of behaviour change communication interventions (in particular, voluntary counselling and testing) [44–46].

To some policy-makers, the slowdown of growth in budgets available for HIV/AIDS programmes is a sobering constraint and makes the potential benefits of radical programmes with high near-term costs irrelevant. Their questions are about the most cost-effective allocation of incremental changes in resources and portfolio optimisation in light of the new data about the additional effect of reducing new infections. To others, the squeeze on funding is a cue to look for ways to drive large reductions in the need for resources in the future, which could be generated by an overhaul of the current epidemic response and an increase in resources in the short term. New, large investments in controlling HIV may not be impossible, but there would have to be a strong case for the return on such an investment.

Estimating Costs

In this collection, Meyer-Rath and Over [19] outline economic concepts that should guide discussions about the potential for ART to reduce incidence, and how the programmatic targets identified by epidemiological modelling could translate into costs. They argue that the nature of the cost function for ART—that is, the cost of providing additional patient-years of ART given the current scale of a programme and practical constraints—has received insufficient attention in earlier analyses. In particular, they suggest that the scale and scope of a country's ART programme, including clinic size and density, cohort maturity, patient mix, and health-worker effectiveness, could mean that the cost of scale-up of ambitious treatment programmes has been substantially underestimated. However, some projected increases in cost could be offset if future programmes radically change by simplifying the delivery of treatment, such as by eliminating measurement of CD4 counts and/or pre-ART disease monitoring. In a commentary in this collection on the review by Meyer-Rath and Over, Bärnighausen et al. [20] consider the dilemma for those making economic projections for the use of ART as prevention. As

ART for prevention represents a significant departure from the current practice of ART for those in greatest clinical need, it is not clear how to extrapolate from past experience.

Short-Term versus Long-Term Goals

Another key economic consideration is finding the right level of spending now to provide the potential for significant benefits in the long term. The HIV/AIDS epidemic and interventions to stop HIV and AIDS are inevitably long-wave events [47], so this issue is particularly important. Treatment programmes spend money today for returns (in terms of averted infections and deaths, or reduced costs) years and even decades later. If future costs and benefits have the same value as current ones, then enormous sums spent today to eventually avert greater costs and reduce mortality forever (if incidence is reduced to low levels) would be judged as a worthwhile expenditure. Whereas, if we accept that, to a decision-maker, savings that are accrued in the future may be worth less than those made today, then potential future payoffs may be less attractive, and investment in programmes for other, more immediate causes of mortality would be a rational, if not necessarily an inspiring or ethical, response. Recognition of this reality for decision-makers requires modellers to vary the relative value that is assigned to costs and benefits in the future. This is called discounting. Discounting is just one component of how decisions are made, and it is important for those contributing to the debate to be able to couch their arguments in the context of this fundamental consideration.

Working within Economic Constraints while Increasing Access

Those making decisions about expanding the provision of ART for prevention purposes must also plan for the long-term maintenance of such a commitment. Once treatment is initiated, it is lifelong, and, because relying on treatment to reduce incidence does not inherently alter the underlying drivers of infectious spread (e.g., patterns of sexual contact), future reduction of an ART intervention effort could lead to a resurgence of the epidemic. Given economic constraints, the most likely scenario might be for programmes to increase access to treatment gradually. They could increase access by expanding eligibility criteria incrementally to include groups who are most likely to benefit clinically, and whose treatment will most reduce onward transmission. Several possibilities for doing this have been raised, including prioritisation according to biological characteristics (e.g., pregnant women, those with active tuberculosis, or those with high plasma viral loads) or according to behaviours (those in serodiscordant couples, those attending sexually transmitted infection clinics, those with many sexual partners, or sex workers). The epidemiological benefit of providing increased access to treatment for groups beyond current guidelines will be determined by the extent to which the criteria being used to prioritise individuals can reliably identify those who most need treatment or contribute most to generating new infections.

There are also many other factors that should be considered in prioritising groups for expanded ART. These include the size of the group and affordability. The cost to access the group is another factor. For instance, would it be less costly to reach pregnant women, who are already in contact with the health system, than some other groups? The response of a group to treatment also needs to be considered. For example, would stable couples adhere to treatment better than others, or would adherence be low if there is little immediate therapeutic benefit? Ethical considerations, programme acceptance, and feasibility also need to be taken into account. For instance, would it be acceptable to provide serodiscordant couples with preferential access to treatment? These layers of considerations will not always point to one particular group as the best option, and local epidemic, economic, and social conditions will also influence this choice. In addition to these judgments being unlikely to be clear-cut, they are further complicated in instances in which human rights and public health do not necessarily have the same objectives if followed to their logical ends, for instance, if the best strategy for a population does not give optimal outcomes for all individuals. In this collection, Delva et al. [21] review these issues for a wide set of prioritisation options, and Boily et al. [22] describe how mathematical modelling can be used to design, conduct, and analyse studies so that the impact of some of these options can be tested and compared effectively.

Full accounting of the economic costs and benefits of ART includes potentially significant macroeconomic benefits (development of infrastructure, supply chains, and education, and productivity gains) and social benefits (reduced orphaning and increased family stability and employment) derived from spending on ART programmes, which could also have synergies with, and spillover benefits for, interventions for other diseases [48]. These important economic questions are not addressed in this collection of articles. Nevertheless, incorporating these factors into estimates of the cost-effectiveness of alternative forms of interventions [49] or estimates of optimal resource allocation [50] among the repertoire of antiretroviral-drug-based and non-antiretroviral-drug-based prevention interventions, even while uncertainties remain, is an important area of ongoing and future research to help inform decision-making processes.

Research Agenda: Upcoming Trials

The findings of the HPTN 052 trial [11] demonstrated the biological efficacy of treatment in reducing infectiousness in heterosexual individuals who receive the best care and monitoring that is possible. The durability of the effect over the long term will be the focus of the next phase of HPTN 052 [51]. The efficacy of ART in reducing infectiousness from anal sex among MSM is being investigated in observational studies, such as the Opposites Attract study in Sydney, Australia (A. Grulich, personal communication).

Meanwhile, the operational questions will centre on how to deliver the services that are required for maximising the impact of treatment on epidemic spread: very high coverage of HIV testing, frequently repeated HIV testing, strong linkage to care, and high retention in care. Many studies that are already underway aim to examine some of these issues [52].

Several large cluster randomized controlled trials that aim to measure the impact of treatment interventions on HIV incidence in whole communities will also be initiated shortly. One of these studies, PopART (HPTN 071) [52], will test the hypothesis that greatly expanded access to treatment, in combination with access to other services including safe medical male circumcision, is feasible and reduces HIV incidence in populations by 60%. The trials will provide an important and direct test of the predictions set out by mathematical models, and models will have a key a role in the design of the studies and the interpretation of findings. In this collection, Boily et al. [22] describe PopART and other upcoming trials, and outline the role of modelling before (in planning and design), during (in monitoring), and after (for interpretation and extrapolation) trials.

Future Directions: Priorities for Modelling

From consultation with programme leaders, key stakeholders, community members, and funders at the HIV Modelling Consor-

Key Points

- It has been established that ART for those infected with HIV can prevent onward transmission of infection, but biological efficacy alone is not enough to confirm the impact that ART could have on the HIV epidemic, or to show how best to use ART to reduce incidence of HIV. This will be among the most important issues in the field of HIV prevention for the foreseeable future.
- Epidemiology, economics, demography, statistics, biology, and mathematical modelling will be central in framing key decisions in the optimal use of ART.
- The HIV Modelling Consortium aims to coordinate and promote research across these disciplines, and facilitate communication between researchers and policy-makers. At a collaborative meeting of this consortium in November 2011, several interlocking themes emerged that are discussed in this article and covered in more depth by other articles in this collection.
- Mathematical modelling is used to investigate the potential impact of treatment on HIV incidence. However, because of incomplete information on all the factors that could influence impact, substantial uncertainties will remain. Models should acknowledge those uncertainties and help prioritise data collection where this could strengthen model conclusions.
- The current economic constraints on HIV prevention bring to the fore the role of modelling to help assess the value and cost-effectiveness of ART. Understanding costs and integrating costing and epidemiological models will be key areas of ongoing and future research to help inform decision-making processes. Models are also being used to help design and interpret trials that test hypotheses about the impact of expanded access to treatment on the spread of HIV in communities.
- We hope that this article and others in the collection will provide a solid foundation upon which greater collaborations between disciplines will be formed, so as to better integrate the role of modelling into the wider scientific process and to more clearly articulate the strengths and weaknesses of particular modelling analyses. This approach will ultimately strengthen the support for evidence-based decision-making in HIV programmes.

tium meeting in 2011, several themes have emerged as priority issues for further analysis and modelling. First, models need to focus more specifically on the impact of decisions over short time horizons, which are of greater relevance when justifying expenditure, as well as on long-term impacts. Second, models should estimate the impact that programmes may have already had and the benefits of current policy decisions, rather than only what impact radically different programmes might have in the distant future. Third, models should estimate impact in a greater variety of settings, including concentrated epidemics. Fourth, models should become better aligned with the experience of real programmes, rather than using unrealistically optimistic assumptions. And, finally, models should be utilised to further explore potential negative outcomes of expanded treatment programmes, such as imperfect adherence, drug resistance, difficulties in recognising resistance if treatment is initiated at high CD4 counts, and the potential influence of compensatory changes in risk behaviours.

At the HIV Modelling Consortium meeting, many called for modelling to articulate the consequences of reductions in funding, such as numbers of new treatment initiations decreasing substantially or even current cohorts of treated patients not being maintained. This modelling would highlight the ethical choices at the heart of these issues. It was also agreed that models should incorporate more fully the benefits of earlier treatment initiation, in terms of a potential, but not certain, additional therapeutic benefit, reduced incidence of tuberculosis, reduced costs of prevention of mother-to-child transmission services, reduced cost of monitoring patients in care and treating opportunistic infections, and spillover effects such as greater productivity and reduced numbers of orphans.

There has also been a call to improve communication of mathematical modelling research to policy-makers, clinicians, and other researchers so as to better integrate its role into the wider scientific process and to more clearly articulate the strengths and weaknesses of particular modelling analyses. In response, Delva et al. [23] in this collection present some principles of "best practice" for model presentation and interpretation, which they hope will become a shared resource for both those who conduct modelling research and those who use modelling results.

Conclusions

The question of how to best use the tools that have been shown to reduce HIV transmission will likely dominate the field of HIV prevention for the foreseeable future. It touches every other aspect of the response to the worldwide HIV epidemics, from the optimal allocation of resources in real programmes, to the relative value of investing in developing additional prevention modalities, to the global spending that will be required in the future. Epidemiology, economics, demography, statistics, and mathematical modelling will be central, and it is hoped that this collection of articles will provide a solid foundation upon which greater collaborations and deeper insights will be formed, and will strengthen the support for evidence-based decision-making, to the benefit of all those whose lives are threatened by HIV epidemics.

Acknowledgments

The following people attended the meeting and meet International Committee of Medical Journal Editors authorship criteria for this article: Till Bärnighausen, Harvard School of Public Health, and Africa Centre for Health and Population Studies; Stephen Becker, Bill & Melinda Gates Foundation; Eran Bendavid, Stanford University; Anna Bershteyn, Intellectual Ventures; John Blandford, US Centers for Disease Control and Prevention; Marie-Claude Boily, Imperial College London; David Burns, US National Institute of Allergy and Infectious Diseases; Valentina Cambiano, University College London; Myron S. Cohen, University of North Carolina at Chapel Hill; Íde Cremin, Imperial College London; Wim Delva, South African Department of Science and Technology/National Research Foundation Centre of Excellence in Epidemiological Modelling and Analysis, University of Stellenbosch, and Ghent University; Christopher Dye, World Health Organization; Jeffrey W. Eaton, Imperial College London; Matthias Egger, University of Bern; Christophe Fraser, Imperial College London; Noya Galai, Johns Hopkins University; Geoff Garnett, Bill & Melinda Gates Foundation; Peter D. Ghys, Joint United Nations Programme on HIV/AIDS; Timothy B. Hallett, Imperial College London; Laura Heaton, US Census Bureau; Charles B. Holmes, Office of the US Global AIDS Coordinator; Jan Hontelez, Erasmus University; Britta Jewell, Imperial College London; Olivia Keiser, University of Bern; Daniel Klein, Intellectual Ventures; Viviane Lima, University of British Columbia; Elisa Long, Yale University; Rob Lyerla, Office of the US Global AIDS Coordinator; Christiaan Marais, HEXOR; Fei Meng, Hasselt University; Gesine Meyer-Rath, Boston University; William C. Miller, University of North Carolina at Chapel Hill; Nicholas Muraguri, Kenya National AIDS & STI Control Programme; Brooke E. Nichols, Erasmus Medical Center; Karima R. Nigmatulina, Intellectual Ventures; Mead Over, Center for Global Development; Nancy Padian, Office of the

US Global AIDS Coordinator; Andrew Phillips, University College London; Carel Pretorius, Futures Institute; Christine Rousseau, Bill & Melinda Gates Foundation; Joshua Salomon, Harvard School of Public Health; Nalinee Sangrujee, US Centers for Disease Control and Prevention; David Serwadda, Makerere University; Frank Tanser, Africa Centre for Health and Population Studies; Juan F. Vesga, Imperial College London; Peter Vickerman, London School of Hygiene & Tropical Medicine; Damian Walker, Bill & Melinda Gates Foundation; Rui Wang, Harvard School of Public Health, and Brigham and Women's Hospital; Alex Welte, South African Department of Science and Technology/ National Research Foundation Centre of Excellence in Epidemiological Modelling and Analysis, and University of Stellenbosch; Richard G. White, London School of Hygiene & Tropical Medicine; Brian Williams, South African Centre for Epidemiology Modelling and Analysis; David P. Wilson, University of New South Wales; David Wilson, World Bank; Basia Zaba, London School of Hygiene & Tropical Medicine.

We also acknowledge David van de Vijver (Erasmus Medical Centre) and Chris Collins (amfAR) for helpful suggestions, and all our collaborators who participated in the HIV Modelling Consortium meeting, "The Potential Impact of Treatment on HIV Incidence"; Stellenbosch, South Africa; 3–5 November 2011. For further details and a list of meeting participants, please see http://www.hivmodelling.org/projects/potential-impact-treatment-hiv-incidence.

Author Contributions

Wrote the first draft of the manuscript: TH. Contributed to the writing of the manuscript: All Members of the HIV Modelling Consortium Treatment as Prevention Editorial Writing Group. ICMJE criteria for authorship read and met: All Members of the HIV Modelling Consortium Treatment as Prevention Editorial Writing Group. Agree with manuscript results and conclusions: All Members of the HIV Modelling Consortium Treatment as Prevention Editorial Writing Group. Contributed to constructing content of this article through 3 day workshop: All Members of the HIV Modelling Consortium Treatment as Prevention Editorial Writing Group.

References

1. Joint United Nations Programme on HIV/AIDS (2011) World AIDS day report: how to get to zero: faster. smarter. better. Available: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216_WorldAIDSday_report_2011_en.pdf. Accessed 25 November 2011.
2. World Health Organization, United Nations Children's Fund, Joint United Nations Programme on HIV/AIDS (2011) Progress report 2011: global HIV/AIDS response. Epidemic update and health sector progress towards universal access. Geneva: World Health Organization.
3. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, et al. (2009) South African national prevalence, incidence, behaviour and communication survey, 2008. A turning tide among teenagers? Cape Town: HSRC Press.
4. Schwartlander B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041.
5. US President's Emergency Plan for AIDS Relief (2011 August) Guidance for the prevention of sexually transmitted hiv infections. Available: http://www.pepfar.gov/documents/organization/171303.pdf. Accessed 25 November 2011.
6. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, et al. (2010) Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science 329: 1168–1174.
7. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, et al. (2010) Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 363: 2587–2599.
8. University of Washington International Clinical Research Center (2011) Pivotal study finds that HIV medications are highly effective as prophylaxis against HIV infection in men and women in Africa. Available: http://depts.washington.edu/uwicrc/research/studies/files/PrEP_PressRelease-UW_13Jul2011.pdf. Accessed 25 May 2012.
9. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, et al. (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med 361: 2209–2220.
10. Celum C, Baeten JM (2012) Tenofovir-based pre-exposure prophylaxis for HIV prevention: evolving evidence. Curr Opin Infect Dis 25: 51–57.
11. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
12. Cohen J (2011) HIV treatment as prevention. Science 334: 1628.
13. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, et al. (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet 375: 2092–2098.
14. Attia S, Egger M, Muller M, Zwahlen M, Low N (2009) Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. AIDS 23: 1397–1404.
15. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
16. Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS (2012) HIV treatment as prevention: the utility and limitations of ecological observation. PLoS Med 9: e1001260. doi:10.1371/journal.pmed.1001260
17. Wilson DP (2012) HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med 9: e1001231. doi:10.1371/journal.pmed.1001231
18. Cohen MS, Dye C, Fraser C, Miller WC, Powers KA, et al. (2012) HIV treatment as prevention: debate and commentary—will early infection compromise treatment-as-prevention strategies? PLoS Med 9: e1001232. doi:10.1371/journal.pmed.1001232
19. Meyer-Rath G, Over M (2012) HIV treatment as prevention: modelling the cost of antiretroviral treatment—state of the art and future directions. PLoS Med 9: e1001247. doi:10.1371/journal.pmed.1001247
20. Bärnighausen T, Salomon JA, Sangrujee N (2012) HIV treatment as prevention: issues in economic evaluation. PLoS Med 9: e1001263. doi:10.1371/journal.pmed.1001263
21. Delva W, Eaton JW, Meng F, Fraser C, White RG, et al. (2012) HIV treatment as prevention: optimising the impact of expanded HIV treatment programmes. PLoS Med 9: e1001258. doi:10.1371/journal.pmed.1001258
22. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250
23. Delva W, Wilson DP, Abu-Raddad L, Gorgens M, Wilson D, et al. (2012) HIV treatment as prevention: principles of good HIV epidemiology modelling for public health decision-making in all modes of prevention and evaluation. PLoS Med 9: e1001239. doi:10.1371/journal.pmed.1001239
24. Garnett GP, Anderson RM (1993) Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. Philos Trans R Soc Lond B Biol Sci 342: 137–159.
25. Eligibility for ART in Lower Income Countries (eART-linc) Collaboration (2008) Duration from seroconversion to eligibility for antiretroviral therapy and from ART eligibility to death in adult HIV-infected patients from low and middle-income countries: collaborative analysis of prospective studies. Sex Transm Infect 84 (Suppl 1): i31–i36.
26. Lodi S, Phillips A, Touloumi G, Geskus R, Meyer L, et al. (2011) Time from human immunodeficiency virus seroconversion to reaching CD4+ cell count thresholds <200, <350, and <500 cells/mm^3: assessment of need following changes in treatment guidelines. Clin Infect Dis 53: 817–825.
27. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
28. Sweat M, Morin S, Celentano D, Mulawa M, Singh B, et al. (2011) Community-based intervention to increase HIV testing and case detection in people aged 16–32 years in Tanzania, Zimbabwe, and Thailand (NIMH Project Accept, HPTN 043): a randomised study. Lancet Infect Dis 11: 525–532.
29. Rosen S, Fox MP (2011) Retention in HIV care between testing and treatment in sub-Saharan Africa: a systematic review. PLoS Med 8: e1001056. doi:10.1371/journal.pmed.1001056
30. Katz IT, Essien T, Marinda ET, Gray GE, Bangsberg DR, et al. (2011) Antiretroviral therapy refusal among newly diagnosed HIV-infected adults. AIDS 25: 2177–2181.
31. Schöni-Affolter F, Keiser O, Mwango A, Stringer J, Ledergerber B, et al. (2011) Estimating loss to follow-up in HIV-infected patients on antiretroviral therapy: the effect of the competing risk of death in Zambia and Switzerland. PLoS ONE 6: e27919. doi:10.1371/journal.pone.0027919
32. Cornell M, Grimsrud A, Fairall L, Fox MP, van Cutsem G, et al. (2010) Temporal changes in programme outcomes among adult patients initiating antiretroviral therapy across South Africa, 2002–2007. AIDS 24: 2263–2270.
33. Brinkhof MW, Dabis F, Myer L, Bangsberg DR, Boulle A, et al. (2008) Early loss of HIV-infected patients on potent antiretroviral therapy programmes in lower-income countries. Bull World Health Organ 86: 559–567.
34. Estill J, Aubriere C, Egger M, Johnson L, Wood R, et al. (2012) Viral load monitoring of antiretroviral therapy, cohort viral load and HIV transmission in Southern Africa: a mathematical modelling analysis. AIDS. E-pub ahead of print. doi:10.1097/QAD.0b013e3283536988
35. Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, et al. (2008) Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. Lancet 371: 1443–1451. doi:10.1016/S0140-6736(08)60624-8
36. Walensky RP, Wood R, Ciaranello AL, Paltiel AD, Lorenzana SB, et al. (2010) Scaling up the 2010 World Health Organization HIV treatment guidelines in

resource-limited settings: a model-based analysis. PLoS Med 7: e1000382. doi:10.1371/journal.pmed.1000382
37. Phillips AN, Pillay D, Garnett G, Bennett D, Vitoria M, et al. (2011) Effect on transmission of HIV-1 resistance of timing of implementation of viral load monitoring to determine switches from first to second line antiretroviral regimens in resource-limited settings. AIDS 25: 843–850.
38. Montaner JS, Lima VD, Barrios R, Yip B, Wood E, et al. (2010) Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 376: 532–539.
39. Hollingsworth TD, Anderson RM, Fraser C (2008) HIV-1 transmission, by stage of infection. J Infect Dis 198: 687–693.
40. Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, et al. (2011) The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet 378: 256–268.
41. Kaiser Family Foundation (2011) International AIDS assistance from donor governments: commitments & disbursements, 2002–2010. Available: http://facts.kff.org/chart.aspx?ch = 946. Accessed 25 November 2011.
42. Menzies NA, Berruti AA, Berzon R, Filler S, Ferris R, et al. (2011) The cost of providing comprehensive HIV treatment in PEPFAR-supported programs. AIDS 25: 1753–1760.
43. Njeuhmeli E, Forsythe S, Reed J, Opuni M, Bollinger L, et al. (2011) Voluntary medical male circumcision: modeling the impact and cost of expanding male circumcision for HIV prevention in eastern and southern Africa. PLoS Med 8: e1001132. doi:10.1371/journal.pmed.1001132
44. Sweat M, Gregorich S, Sangiwa G, Furlonge C, Balmer D, et al. (2000) Cost-effectiveness of voluntary HIV-1 counselling and testing in reducing sexual transmission of HIV-1 in Kenya and Tanzania. Lancet 356: 113–121.
45. Allen S, Serufilira A, Bogaerts J, Van de Perre P, Nsengumuremyi F, et al. (1992) Confidential HIV testing and condom promotion in Africa. Impact on HIV and gonorrhea rates. JAMA 268: 3338–3343.
46. Nuwaha F, Tumwesigye E, Kasasa S, Asiimwe S, Wana G, et al. (2009) Population-level changes in knowledge of HIV status, stigma and HIV risk behaviors after district-wide door to door voluntary counseling and testing (VCT) in Bushenyi District, Uganda [abstract]. 16th Conference on Retroviruses an Opportunisitic Infections; 8–11 February 2009; Montreal, Canada.
47. The aids2031 Consortium (2011) AIDS: taking a long-term view. Upper Saddle River (Jew Jersey): FT Press Science.
48. Resch S, Korenromp E, Stover J, Blakley M, Krubiner C, et al. (2011) Economic returns to investment in AIDS treatment in low and middle income countries. PLoS ONE 6: e25310. doi:10.1371/journal.pone.0025310
49. Tan-Torres Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, et al (2004) Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization.
50. Brandeau ML, Zaric GS, Richter A (2003) Resource allocation for control of infectious diseases in multiple independent populations: beyond cost-effectiveness analysis. J Health Econ 22: 575–598. doi:10.1016/S0167-6296(03)00043-2
51. Cohen MS, Muessig KE, Smith MK, Powers K, Kashuba AD (2012) Antiviral agents and HIV prevention: controversies, conflicts and consensus. AIDS 12. E-pub ahead of print. doi:10.1097/QAD.0b013e3283543e83
52. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res. 9: 446–469.
53. Abu-Raddad LJ, Longini IM Jr (2008) No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. AIDS 22: 1055–1061.
54. Eaton JW, Hallett TB, Garnett GP (2011) Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS Behav 15: 687–692.
55. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, et al. (2012) Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. AIDS Behav 16: 312–322.
56. Todd J, Glynn JR, Marston M, Lutalo T, Biraro S, et al. (2007) Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. AIDS 21(Suppl 6): S55–S63.
57. Reynolds SJ, Makumbi F, Nakigozi G, Kagaayi J, Gray RH, et al. (2011) HIV-1 transmission among HIV-1 discordant couples before and after the introduction of antiretroviral therapy. AIDS 25: 473–477. doi:10.1097/QAD.0b013e3283437c2b
58. World Health Organization (2010) Antiretroviral therapy for HIV infection in adults and adolescents. Recommendations for a public health approach. 2010 revision. Geneva: World Health Organization.
59. World Health Organization (2012 April) Guidance on couples HIV testing and counseling including antiretroviral therapy for treatment and prevention in serodiscordant couples. Geneva: World Health Organization.
60. World Health Organization (2011) WHO consultation on the strategic use of antiretrovirals for treatment and prevention of HIV. Geneva: World Health Organization.

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Modelling the Cost of Antiretroviral Treatment—State of the Art and Future Directions

Gesine Meyer-Rath[1,2]*, Mead Over[3]

1 Center for Global Health and Development, Boston University, Boston, Massachusetts, United States of America, **2** Health Economics and Epidemiology Research Office, Department of Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, **3** Center for Global Development, Washington, District of Columbia, United States of America

Abstract: Policy discussions about the feasibility of massively scaling up antiretroviral therapy (ART) to reduce HIV transmission and incidence hinge on accurately projecting the cost of such scale-up in comparison to the benefits from reduced HIV incidence and mortality. We review the available literature on modelled estimates of the cost of providing ART to different populations around the world, and suggest alternative methods of characterising cost when modelling several decades into the future. In past economic analyses of ART provision, costs were often assumed to vary by disease stage and treatment regimen, but for treatment as prevention, in particular, most analyses assume a uniform cost per patient. This approach disregards variables that can affect unit cost, such as differences in factor prices (i.e., the prices of supplies and services) and the scale and scope of operations (i.e., the sizes and types of facilities providing ART). We discuss several of these variables, and then present a worked example of a flexible cost function used to determine the effect of scale on the cost of a proposed scale-up of treatment as prevention in South Africa. Adjusting previously estimated costs of universal testing and treatment in South Africa for diseconomies of small scale, i.e., more patients being treated in smaller facilities, adds 42% to the expected future cost of the intervention.

Introduction

Informed by biological plausibility [1], observational studies [2], and a trial [3] showing that ART reduces transmission of HIV within heterosexual serodiscordant couples, recent modelling papers [4–6] have projected the reduction in HIV incidence and the impact on health care costs that would follow from achieving close-to-universal coverage with HIV testing and ART. These papers argue that sufficiently universal ART coverage would eventually pay for itself by suppressing HIV incidence and therefore averting the future need for HIV care, including ART. Other papers in the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections" analyse the sensitivity of the projected population-level incidence reductions to the structure and assumptions of an epidemiological projection model [7–9]. This paper focuses on the cost side of such projection models. We begin with a general discussion of cost accounting identities versus flexible cost functions. Then we review the available literature on modelled estimates of the projected cost of ART provision, including ART for prevention, with a focus on identifying determinants authors have included, implicitly or explicitly, in their assumed cost function for ART service delivery. We then discuss the evidence for a number of such cost determinants. Finally, we present an example of a flexible cost function used to explore how economies of scale might affect the costs of scaling up ART in South Africa. A second paper focussing on economic evaluation in this collection further discusses how operational and effectiveness issues in scaling up ART for prevention will affect its cost-effectiveness [10].

Review articles synthesize in narrative form the best available evidence on a topic.

Cost Accounting Identities versus Flexible Cost Functions

Just as most epidemiological projection models include a functional representation of epidemiological concepts such as the force of infection, cost projection models include a function or a set of functions to characterise the relationship between the total cost of ART service delivery and various determinants of cost, such as the number of patients on treatment, the stage in their disease at which they were recruited, and the ART regimen they receive. Most existing cost projections assume a single constant unit cost per patient-year, or per patient-year on a certain regimen, across large populations and often extended projection periods. A somewhat more complex approach is to assume a single unit cost for each of a set of services received by an HIV-positive patient, such as a unit cost for each type of laboratory test or outpatient

Citation: Meyer-Rath G, Over M (2012) HIV Treatment as Prevention: Modelling the Cost of Antiretroviral Treatment—State of the Art and Future Directions. PLoS Med 9(7): e1001247. doi:10.1371/journal.pmed.1001247

Academic Editor: David R. Bangsberg, Massachusetts General Hospital, Harvard Medical School, United States of America

Published July 10, 2012

Funding: GMR was supported by the HIV Modelling Consortium (funded by the Bill & Melinda Gates Foundation) for work towards this paper. Support from the Bill & Melinda Gates Foundation to the Center for Global Development also finances MO's work on the economics of AIDS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: ART, antiretroviral therapy; LMIC, low- and middle-income country; NGO, non-governmental organisation.

* E-mail: gesine@bu.edu

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

visit or inpatient day, and then multiply these unit costs by an estimate of the number of each of these services per patient-year and by the number of patient-years delivered in a year. We call such an equation an accounting identity and designate a total annual cost so defined as an accounting identity cost function, TC_{AI}. In its simplest form such a cost function can be written as

$$TC_{AI} = \text{Fixed cost} + A\sum_{k} q_k \quad (1)$$

where k indexes the facilities delivering ART, q_k represents the output of facility k in a single time period, typically a year, and A is the average variable cost per patient-year. Cost accounting identities impose the discipline of arithmetical consistency on discussions of budgets, costs, expenditures, and efficiency, and predict future expenditures over the short run. They are a natural approach when estimating how much delivery of a service "should" cost. (See the discussion of the distinction between "normative" and "positive" cost functions in Text S2.) As such, they are often sufficient for capturing the impact of incremental policies, e.g., an extension of a health care intervention to a slightly larger proportion of the same population by increasing coverage.

However, cost accounting identities cannot be used to predict how costs will change when civil servants, managers, providers, and patients have an opportunity to adjust service delivery by, for example, substituting one input for another, or changing the scale and scope of operations, eligibility criteria, task shifting, or the deployment of supply- or demand-side incentives. We argue that, as a result, cost accounting identities are too rigid to model large-scale changes over periods of more than a few years—such as those required to achieve the HIV prevention benefits of ART. For these purposes, a more flexible cost function such as

$$[\![TC]\!]_F = f(p_i, Z_j, q_k) \quad (2)$$

can provide a more plausible characterisation and projection of total annual costs. In Equation 2, p and Z are vectors representing, respectively, the set of relevant input prices and all other policy and environmental determinants of cost, many of which we discuss in this paper. The notation $f(\ldots)$ stands in for a flexible functional form chosen either to fit the data or, when data is lacking, to fit the analysts' assumptions (see Text S2 for more details). For simplicity, in both Equation 1 and 2 we have suppressed the time subscripts, but in a more formal development, time might itself influence price, output, or other policy determinants.

The Use of Cost Functions in Published Modelled Economic Analyses of ART

In order to determine the current state of the art, we reviewed the available literature on modelled estimates of the projected cost of ART provision to a variety of eligible populations, including ART for prevention. We searched eight databases (PubMed, HealthSTAR, POPLINE, EconLit, HEED, Web of Knowledge [Science and Social Sciences], Embase and CAB Health) for the years 1988–2011 using any combination of the terms *cost**, *econ**, and *HIV* or *AIDS*. We supplemented the identified articles by reviewing the reference lists of identified articles, additional review articles, and grey literature (slides, conference proceedings, books, and manuals). We included all articles in any language that contained modelled cost data of any kind as well as ART as an intervention, except where it was used for the prevention of mother-to-child transmission only. Abstracts and articles in all languages (English, Italian, Spanish, French, and German) were read in full by the first author, who made the decision whether to include the article in the review. We excluded editorials and letters, articles without quantitative data, and articles that did not include a modelled estimate, such as papers reporting cost data from a single site. The last have been reviewed repeatedly in the past [11–15]. We reviewed the included articles with regards to their economic evaluation method, the type of model used, their time horizon, the outcome metric and result, and whether the input cost (often in the form of average per patient cost per unit time) was constant or had been varied by determinants such as types of regimens used, health state, time on treatment, and mode of delivery, in either the main or the sensitivity analysis.

We identified 45 published articles, one conference abstract, and four reports on modelled economic analyses of ART provision worldwide (Table 1; Text S1). Thirty-eight analyses were for single countries, four were for wider regions, and eight were global. Five analyses, all for single countries, specifically considered the impact of ART on HIV transmission; we discuss these separately.

Thirty-three analyses modelled ART programmes within a single country, without considering the transmission impact of ART [16–48]. Most of the 24 high-income-country analyses compared the incremental cost and effectiveness of a new drug regimen with that of an older one [22,24–26,33,36–39]. Amongst the nine low- and middle-income-country (LMIC) analyses, six analyses focussed on the choice of eligibility criteria [40–45,47,48]. One analysis compared ART with no ART [43], one, first-line treatment with first- and second-line treatment [40], and one, different regimens for women previously exposed to single-dose nevirapine as part of prevention of mother-to-child transmission [47].

In terms of the use of cost functions, most of these single-country papers varied input cost (i.e., the cost per patient per unit of time) by protocol-related variables such as treatment regimen, health state (defined by the absence or presence of symptoms, opportunistic infections, AIDS-defining diseases, and/or CD4 cell count levels), and/or time on treatment (see Table 1). Only two papers, both of them on LMICs, varied cost by level of care (secondary versus tertiary) [43] or mode of health care provision (public versus private) [44]; none of the papers varied per patient cost by scale or other programmatic variables.

The four regional studies [49–52] all focussed on sub-Saharan Africa (with one study [52] additionally including Southeast Asia). These studies modelled the cost of defined increases in ART coverage from a low baseline [49,50] and the cost effectiveness of ART provision through the specific setting of an antenatal care clinic [51]. One paper used the same constant input cost for all patients [52]; two papers varied input cost by regimen [49,52]. None of the papers varied per patient cost by any other variables.

The eight global studies, published between 1997 and 2011, describe a clear evolution in both data availability and modelling technique [53–60]. The older analyses estimate cost based only on the number of HIV-positive people from a number of sources, varying assumptions of ART coverage at baseline, with costs based on guidelines and prices from high-income countries [53,54]. Later analyses model global cost under concrete programmes, such as the World Health Organization's 3 by 5 initiative [57] and the Global Fund to Fight AIDS, Tuberculosis and Malaria [56,58], based on per patient cost estimates from relevant LMICs and more advanced epidemiological models of the number of patients in need of ART, such as the Spectrum model [58,59] and the Resource Needs Model [60]. Three of the eight global analyses used constant input costs for all patients [53–55]; two varied input cost by regimen [57,58], and one additionally by health state [58]. One study included the impact of access to pool procurement

Table 1. Overview of the methods and results of previously published modelled economic analyses of antiretroviral treatment.

Category	Number of papers	Number with No Variation in Input Cost in Main Analysis	Number with Unit Cost Held Constant within Each Determinant				Number with Sensitivity Analysis			Results in 2011 US Dollars		
			Regimen	Health State	Time on Treatment	Other Variables	Done (of Which Probabilistic)	Includes Drug Cost	Includes Other Cost	Cost per Life Year Saved	Cost per QALY Gained	Total Annual Cost
No impact of treatment on transmission assumed												
Single country, high income	24	2	18	23	1	US state: 1; medication payer: 1	20–21[a] (1)	14	13	9,027 –84,882[b]; 13,781 –42,944[c]; cost savings: 71,111[d]	20,885 –40,279[c]; 16,430 –295,113[d]	
Single country, low/middle income	9	1	5	6	2	Inpatient cost by level of care: 1; mode of delivery (public versus private): 1	5 (1) plus 1 scenario analysis	6	4	40 –2,540[d]	1,098 –18,851[d]	
Regional	4	1–2[e]	2	0	0	0	1 (0)	1	1			2.3 billion[d] –9.3 billion[d]
Global	8	3	2	1	0	Access to CHAI prices: 1; GNP per capita: 1; shift to primary care and cheaper diagnostics: 1	4 (1)	2	3			1.8 billion[d,f] –176 billion[d,g]
Impact of treatment on transmission assumed												
Single country, high income	1	0	0	1	0	0	1 (0)	0	0		20,542 –22,649[d]	
Single country, low/middle income	5	0	3	3	1	Unstructured versus structured treatment provision: 1	4 (0)	2	2		Cost savings: 4,594[d]	

[a]One publication ([28]; in abstract format) has no information on whether sensitivity analysis was conducted.
[b]For zidovudine monotherapy.
[c]For dual therapy.
[d]For highly active ART.
[e]One study [50] does not supply enough information on ART input cost to know whether it is constant.
[f]Analysis from 2011, based on country-level cost data.
[g]Analysis from 1997, based on high-income country cost data extrapolated worldwide.
CHAI, Clinton HIV/AIDS Initiative; GNP, gross national product; QALY, quality-adjusted life year.
doi:10.1371/journal.pmed.1001247.t001

prices negotiated by the Clinton HIV/AIDS Initiative on per patient cost [57], one varied drug prices by per capita gross national product [56], and one assumed a reduction of per patient cost of 65% by 2020 as a result of task shifting and cheaper point-of-care diagnostics [60]. No other cost determinants were considered.

Five studies between 2006 and 2011 that analysed the cost of ART for a single country included an impact of treatment on HIV transmission and, hence, on the number of future infections and future cost [4,5,61–63]. Three of these analyses were cost-effectiveness analyses of different strategies of eligibility and coverage [61–63]; two were analyses of the cost impact and cost benefit of earlier treatment initiation, including universal testing and treatment [4,5]. With respect to cost functions, three of the analyses varied input cost by regimen [4,5,63], three by health state [61–63], and one by time on treatment [62]; additionally, one analysis varied input cost by whether treatment was administered in a structured way in the public sector or an unstructured way in the private sector [62]. No other variation in cost was considered.

Potential Determinants of a Flexible Cost Function

As summarised above, most modelled estimates of the projected cost of ART provision to date have used cost accounting identities, with minimal use of cost functions. If a more flexible cost function is chosen for modelling the future cost of ART over several decades, which variables should be included in this function? Here and in Table 2, we review the evidence for some possible determinants of the cost of ART provision.

Treatment Characteristics: Regimens, Health States, Time on Treatment

Most reviewed papers recognised that more complex cases of any disease engender higher treatment costs. Modellers addressed this by assuming a unit cost that varied by treatment regimen, health state, or time on treatment. These are important cost determinants, since the cost of a national programme will be largely defined by the distribution of the national treatment cohort into first- and second-line regimens (with second-line regimens being much more expensive in most countries) [64] and into CD4 cell count strata associated with different disease burden and cost. Likewise, an analysis of hospitalisation frequency and cost in the same patients before and after ART initiation found the cost of hospitalisation per patient-year in patients with CD4 cell count <100 cells/μl to be ten times higher than in patients with CD4 cell count >350 cells/μl [65] (see also [66–68]). However, we argue that these characteristics are not the only ones that input cost should vary by, and their relevance for total cost might be overwhelmed in situations of rapid scale-up or large-scale changes to programme delivery such as task shifting to lower levels of facilities and health care cadres.

Factor Prices

The prices of factors of production, including labour, supplies, utilities, transportation, equipment, and buildings, clearly affect the cost of health services. By varying input cost by treatment regimen and, in some cases, also changing the cost of laboratory tests over time, most of the reviewed analyses have taken factor prices into account. And for good reason: the cost of antiretroviral drugs—in many countries the largest component of the cost of ART provision—has changed dramatically over the last ten years, especially for LMICs. By October 2000, the prices of antiretroviral drugs in resource-constrained settings had fallen by 90% on average [69], owing largely to the increased availability of generically manufactured drugs from three Indian companies and the possibility of importing these drugs in parallel with patent-protected drugs under the World Trade Organization Agreement on Trade-Related Aspects of Intellectual Property Rights [70]. The price of the non-generic version of the most common first-line drug combination (stavudine+lamivudine+nevirapine) dropped by 93% from US$10,439 to US$727 between June 2000 and September 2001 [71]. Even though the price of the regimen fell by another 54% between 2001 and 2008, the scope for further reductions in the price of antiretrovirals is assumed to be limited, shifting the focus to the cost of other factor prices such as service delivery, laboratory tests, and overheads.Reductions in all of these are targeted by UNAIDS's Treatment 2.0 initiative [72].

Scale

As mentioned, none of the reviewed papers considered an impact of scale, i.e., the size or coverage of the programme, on cost, despite the dramatic increases in scale modelled by some of the papers—especially those analysing the cost of treatment for prevention [4,5]. This stands in contrast to much of economic

Table 2. Schematic summary of determinants of the cost of ART provision.

Determinant	Metric	Direction and Size of Change in Cost	Direction of Change with Scale	Open to Direct Manipulation?
Treatment characteristics: regimens, health states, time on treatment	Median CD4 cell count under ART; distribution into first line/second line; proportion of cohort with CD4 <50 cells/μl	↓ ↓	↑	No
Factor prices	Cost per input	↓ / ↑	↓	(Yes)
Scale	Number of patients; number of ART clinics	↓ ↓, then ↑	—	(Yes)
Experience of facility and programme	Total patient-years of treatment	↓	↑	No
Scope (facility type) and distribution into care sectors	Proportion treated in primary- versus secondary-versus tertiary-level clinics versus stand-alone clinics; proportion treated by public versus private (for-profit and not-for-profit)	↓ / ↑	↑	Yes
Quality of care	Retention ± clinical improvement (weight, CD4 cell count, viral load)	↑, then ↓	↓	Yes
Technical efficiency: incentives, supervision, and technical change	Provider payments as a function of output or outcome; frequency/intensity of supervision/ training; doctor/nurse ratio or protocol selection	↓, except technical change:?	?	Yes

doi:10.1371/journal.pmed.1001247.t002

theory, which assumes a U-shaped relationship between scale and average cost, with cost per unit of output at first decreasing as quantities of output increase, because inputs (e.g., staff) are shared to produce an increasing number of outputs (e.g., patients seen). When scaling up further, beyond a certain number of outputs, new inputs will be required, leading to increasing average cost for large facilities or broadly expanded programmes. Scale economies seem plausible in ART service delivery because the cost of some functions of an ART treatment site, such as building maintenance, personnel management, and the transportation of supplies, will increase in more direct proportion to the number of sites than to the number of patients each one serves. This means that at the site level, increasing the number of patients generates a less than proportionate increase in cost.

Only a few programmes have produced data that have allowed this relationship to be examined empirically. Economies of scale have been found in HIV prevention programmes [73–75] and in the modelled cost of hygiene outreach interventions, the latter showing a U-shaped relationship between coverage and average or marginal cost [76]. The worked example below and Text S2 provide more discussion of the concept and application of scale economies.

Experience of Facility and Programme

The implementation of most interventions is traditionally assumed to benefit from "learning by doing", which results in reductions in average cost. Since this learning often coincides with scale-up, this relationship is not always easy to distinguish from the reduction of average cost with scale mentioned above. In an analysis of data from ART clinics supported by the US President's Emergency Plan for AIDS Relief, Menzies et al. found that median per patient cost across a number of sites in different countries decreased with each successive six-month period from the start of the ART programme at each site [77], with the biggest decrease between the first and the second six-month periods. The potential effect on cost of increased facility and programme experience over time was not considered in any of the reviewed papers.

Scope (Facility Type) and Distribution into Care Sectors (Private versus Public)

As with scale, the cost of a national ART programme will also be affected by a change in the scope of ART provision, i.e., the type of facilities (e.g., primary health care clinics versus specialised ART clinics at secondary- or tertiary-level hospitals) and whether or not they are in the public or the private sector, with the private sector further divided into for-profit and not-for-profit (e.g., non-governmental organisations [NGOs]). Generally, larger health care facilities, such as hospitals, can achieve economies of scope by spreading the cost of infrastructure over the production of multiple health services. Rosen et al. compared the cost of ART provision per patient-year for the first 12 months of treatment across a clinic in a public hospital, a group of private general practitioners, a private NGO-run HIV clinic, and a private NGO-run primary health care clinic in South Africa [78]. They found costs to vary significantly between sites as a result of differences in service delivery (see Figure 1). Since patient mix was comparable across three of the four sites, only a small portion of the difference in cost could be ascribed to differences in disease severity. Amongst the reviewed papers, only three included level of care as a variable determining input cost (in South Africa [43], India [44], and Thailand [63]). Future cost projections should include information on the variation of cost by level of care and mode of delivery, as well as the expected distribution of the treatment cohort between different levels and modes, especially where these are likely to change as a result of planned dramatic increases in the size of the programme.

Quality of Care

Quality of health care is notoriously difficult to measure, but in ART service delivery, a facility's success at retaining patients in treatment, and improving the patient cohort's health on average, is a reasonable proxy. The same analysis by Rosen et al. compared the cost per quality-adjusted output between the four settings, using routinely collected data (such as patient status, CD4 cell counts, viral loads, and the absence or presence of new World Health Organization stage 3 or 4 conditions) to calculate patient retention in care and response to treatment [78]. While the cost of patients who were no longer in care (i.e., had died or been lost to follow-up during the first 12 months after treatment initiation) was comparable across settings, the cost per patient in care and responding to treatment, and the cost per patient in care and not responding to treatment, was significantly different between the four clinics (Figure 2). Depending on the quality of care in each clinic, and the resulting levels of loss to follow-up and treatment failure, the additional cost per patient in care and responding was 22% and 48% of the average annual cost per patient at two sites because of resources spent on patients either leaving care or not responding to care.

Technical Efficiency: Incentives, Supervision, and Technical Change

Technical efficiency is defined as the production of a good or service without waste, and is thus another important determinant of cost. Both public and private sector providers face constraints in the availability and quality of staff, which will affect the cost of rolling out an intervention differently at a different scale. Staffing in the public sector faces constraints such as lower wages, low work morale, and staff absenteeism, which result in low quality of care. Staffing in the private sector may not be subject to those issues to the same extent because of fee-for-service financing mechanisms, but fee-for-service mechanisms have the undesirable effect of deterring patients, especially uninsured patients, from seeking treatment [79]. Leonard and colleagues have shown that non-financial incentives such as encouragement and supervision by a peer can improve the quality of care provided by health care workers [80,81]. As donor programmes such as the US President's Emergency Plan for AIDS Relief and its contractors relinquish direct control of patient treatment in favour of subsidies to NGOs or technical support for local government provision, the issue of management will become increasingly important as a determinant of technical efficiency and therefore costs.

Our view that programme characteristics such as scale/coverage, scope, managerial incentives, and quality/effectiveness can have important effects on the costs of ART delivery is endorsed by a second paper in this *PLoS Medicine* collection [10], which also points to the difficulty of projecting the future costs of technologies that are not yet widely used or have not even been invented. The solution to the former problem is to collect cost data on a wide range of current practices, and project future costs under the hypothesis that the technology mix will shift, e.g., towards smaller scale treatment programmes, as in the example in the next section. Projecting the costs of unknown future innovations is a less tractable problem, but arguably could best be approached by using simple flexible functions of a few fundamental variables like input prices, and allowing technical efficiency to improve according to a time trend.

A Worked Example of a Flexible Cost Function: The Impact of Scale on the Cost of Universal Testing and Treatment

For achieving the target coverage for universal testing and treatment in South Africa, Granich et al. [4] proposed a scale-up

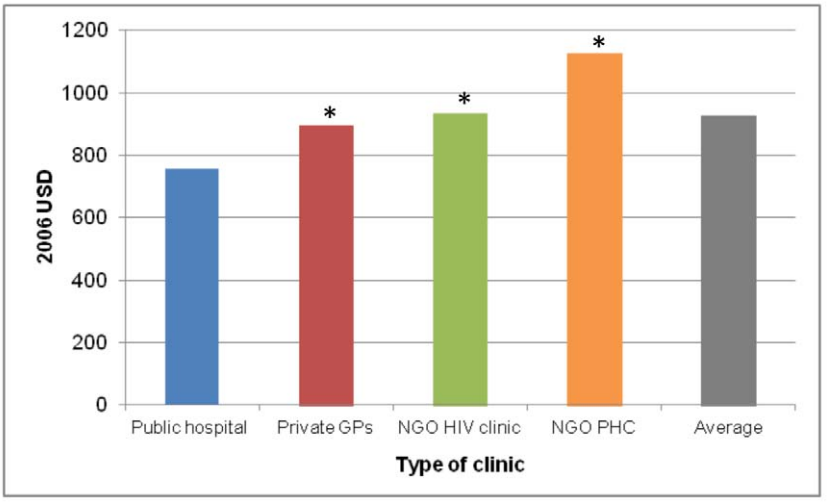

Figure 1. Annual per patient cost of ART provision in four different settings in South Africa. Based on [78]. *, difference from public hospital significant at $p<0.05$. GPs, general practitioners; PHC, primary health care clinic; USD, US dollars.
doi:10.1371/journal.pmed.1001247.g001

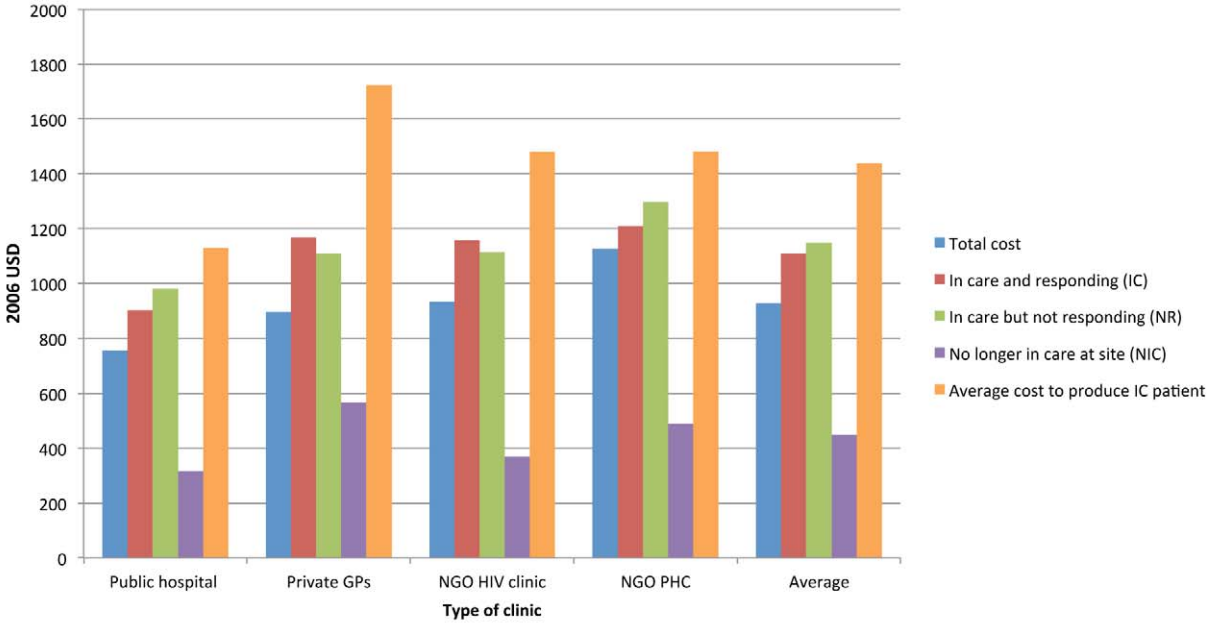

Figure 2. Annual per patient cost of ART provision per type of outcome in four different settings in South Africa. Based on [78]. GPs, general practitioners; PHC, primary health care clinic; USD, US dollars.
doi:10.1371/journal.pmed.1001247.g002

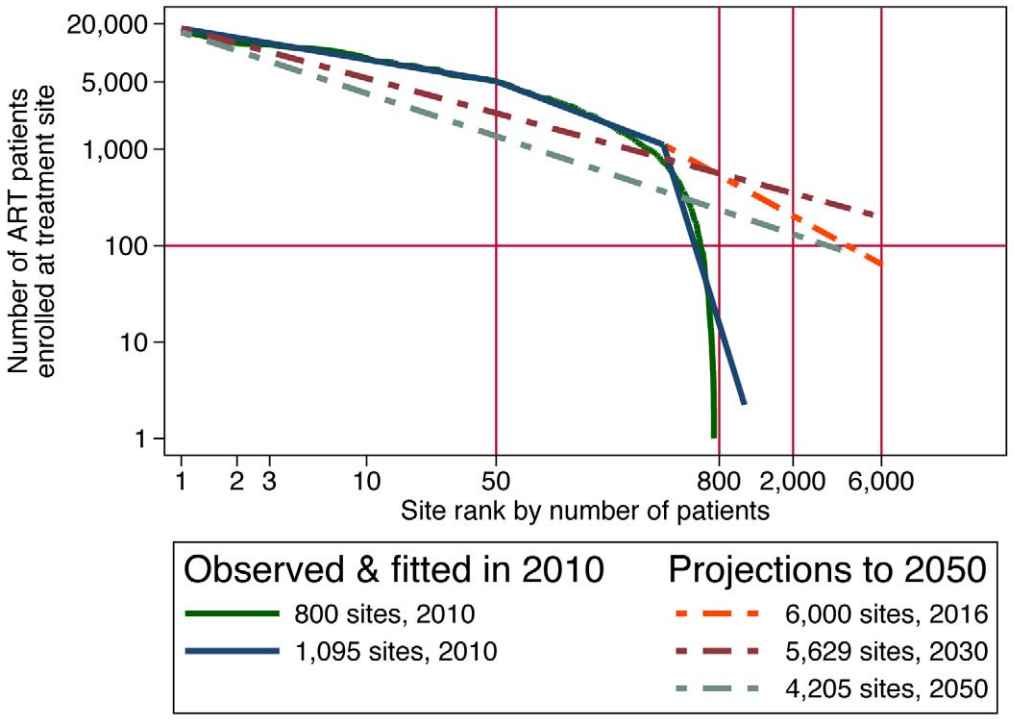

Figure 3. Size-rank distribution of ART facilities in 2010 and projected to future years in order to implement a universal test-and-treat strategy in South Africa.
doi:10.1371/journal.pmed.1001247.g003

from 1.5 million patients on ART in June 2011 [82] to 4.1 million patients by mid-2016. While a flexible cost model of this scale-up proposal could incorporate any of the cost determinants described above, we have data on only one of these: the current size distribution of treatment facilities, i.e., scale. Since economies of scale seem likely to be a persistent feature of ART service delivery, we use this cost determinant in this example, with the hope that more of the data needed to model other potentially important cost determinants will become available in the future. We reviewed the actual size distribution of accredited ART treatment sites in South Africa in June 2010, using government and other sources (Figure 3). When the logarithm of size is charted against the logarithm of size rank, many size distributions in nature are approximately linear, following Zipf's law [83]. We hypothesize that the marked nonlinearity of the size distribution of South Africa's ART sites in 2010 was due to the recent scale-up occurring in larger sites and was temporary. If that is true, and if the largest 50 sites are assumed to retain their current patient loads during programme expansion, then expansion from 1 million patients at the beginning of 2010 to 4.1 million in 2016 would require that more sites be opened and that the scale of smaller sites be increased sufficiently to accommodate the additional patients. As a result, the size distribution of ART sites would straighten out over time. Then, as patient load subsequently contracts over time due to the hypothesized prevention success of the universal test-and-treat policy, we expect the size distribution to mature into a power law that is linear in logarithms, which first steepens, as smaller sites contract first and, once the number of enrolled patients contracts to below 1 million, contracts proportionally at all sites (see Text S2 for details).

Assuming a plausible size distribution of the patient load at ART sites allows us to estimate the effect that a cost function incorporating scale economies would have on the projection of total cost. Suppose that the production technology of ART services exhibits a scale elasticity of 0.7, meaning that every 10% increase in scale is associated with only a 7% increase in total cost, because of scale efficiencies. Assuming for simplicity that all economies of scale occur at the facility level, total cost (tc) for the country would be the sum of

$$\mathrm{tc}_k = A_k * q_k^{0.7} \tag{3}$$

over all the sites in the country, where $A_k = f(p_k, z_k)$, held constant at $\bar{A}$ in the present analysis (US$7,600, calibrated from the known size distribution of patients and total cost per patient in 2010; Text S2 gives results for other elasticities of scale between 1.0 and 0.5). Since average cost at a site is defined as total cost at that site divided by quantity of patients at that site, the facility-specific average cost function (atc_k) consistent with Equation 3 is

$$\mathrm{atc}_k = \frac{\mathrm{tc}_k}{q_k} = \bar{A} * q_k^{-0.3} \tag{4}$$

Applying this cost function to the current and projected facility size distribution yields estimates of the total national cost of ART in each year of the simulation, which we compare to the Granich et al. estimates for the same scale-up scenario (Figure 4). Our assumption that the number of clinics must expand substantially to serve the estimated 4.1 million patients means that an increasing proportion of patients will be served in smaller clinics, which suffer from diseconomies of operating at small scale. In comparison to Granich's estimate of a peak annual cost of US$3.5 billion in 2016, the scale-adjusted estimate is US$4.4 billion, or 26% higher. As the number of patients moderates over time (due to Granich et al.'s

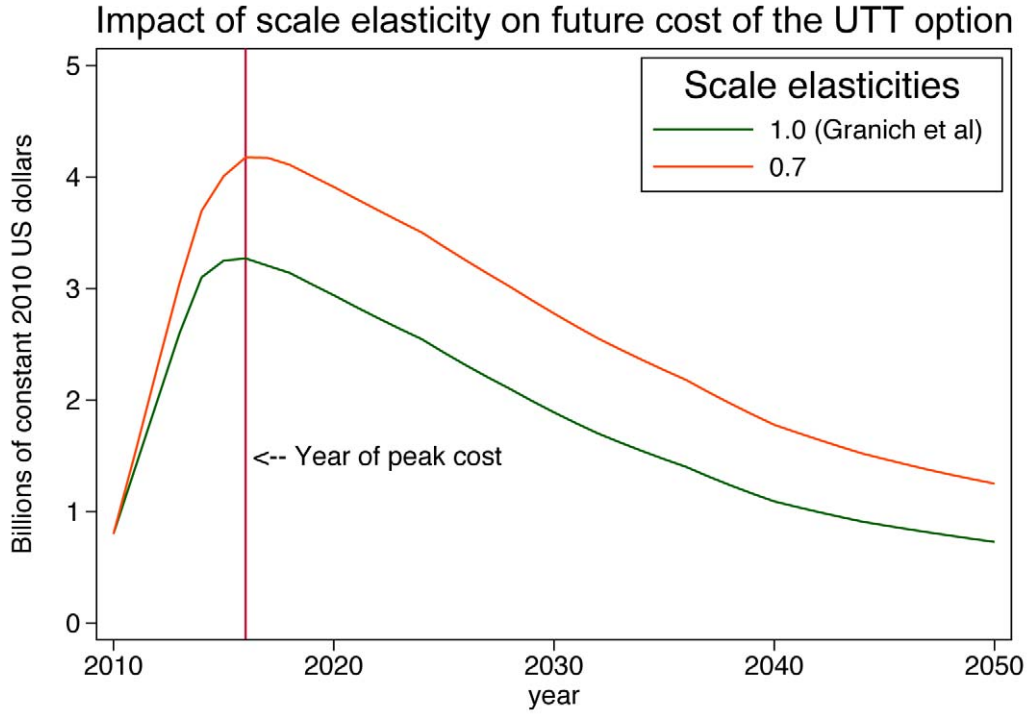

Figure 4. Impact of scale elasticity on future cost of a universal test-and-treat strategy in South Africa. UTT, universal testing and treatment.
doi:10.1371/journal.pmed.1001247.g004

assumptions of a strong population-level reduction in HIV transmission and of a concomitant 40% reduction in risky behaviour), the excess of scale-adjusted costs over accounting identity costs declines to below 20% and then rises again to 50% by the year 2050. Total accumulated cost over the 40-year period of the projection rises from US$75 billion to US$106 billion, an increase of 42%.

This example shows that the simple adjustment of the cost per patient-year for scale and a plausible pattern of distribution of patients into clinics can have a major impact on projected costs over future decades and can highlight the challenge of scaling up a treatment programme to full coverage of people outside urban areas.

Conclusions and Recommendations for Improved Cost Projections

For modellers' projections of alternative ART scale-up scenarios to attract serious policy attention, the assumptions and structure of the cost side of these models, like those on the epidemiological side, should be plausible, supported by observational studies, and, where available, based on results from trials of the costs of alternative service delivery methods. The envisaged cost-effectiveness analyses alongside the planned large-scale trials of treatment as prevention that will be rolled out over the next years provide a historic opportunity to collect such data and allow more precise projections of the future cost of ART programmes using flexible cost functions. Text S2 provides a summary of the differences in data and algebra needed for an accounting identity versus a flexible cost function for estimating cost for an individual facility's or a country's national ART programme. Data collection on large samples of facilities should go beyond measuring the quantity and quality of ART services, to capturing the actual cost of services delivered in a sample of facilities at different levels of care and details about all of the above-listed determinants of cost. With such data on a sample of ART facilities within its own borders, a

Key Points

- In modelling the projected costs of a health programme, flexible cost functions, in which costs vary with certain known or assumed determinants, provide a more plausible characterisation and projection of total annual costs than simple accounting identities.
- A review of previous models estimating the cost of ART provision indicates that while most models accounted for how costs vary with patient health status and treatment regimen, variability in other determinants of cost was rarely included.
- Potential determinants of cost that could be included in flexible cost functions for ART provision when modelling over several decades into the future include patient health status and treatment regimen, factor prices, programme/facility scale, facility experience, facility type, quality of care, and the technical efficiency of staff.
- A worked example of a flexible cost function modelling the impact of one of these determinants, programme scale, on the costs of a proposed universal testing and treatment programme in South Africa found that the inefficiencies of small scale could add up to 42% to the total future cost of the programme.
- Another article in this *PLoS Medicine* collection [10] discusses additional operational and effectiveness issues relevant for the economic evaluation of scaling up ART for prevention.

country's government and any donors supporting its HIV care programme can not only improve their projections of the long-term implications of any given commitment to antiretroviral treatment, but also model the benefits of policies to improve the cost-effectiveness of their efforts.

Supporting Information

Text S1 Methods and results of previously published modelled economic analyses.
(PDF)

Text S2 The algebra of flexible cost functions.
(PDF)

References

1. Gilliam BL, Dyer JR, Fiscus SA, Marcus C, Zhou S, et al. (1997) Effects of reverse transcriptase inhibitor therapy on HIV-1 viral burden in semen. J Acquir Immune Defic Syndr 15: 54–60.
2. Montaner JSG, Lima VD, Barrios R, Yip B, Wood E, et al. (2010) Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 376: 532–539.
3. Cohen M, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
4. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
5. Hontelez JAC, de Vlas SJ, Tanser F, Bakker R, Bärnighausen T, et al. (2011) The impact of the new WHO antiretroviral treatment guidelines on HIV epidemic dynamics and cost in South Africa. PLoS ONE 6: e21919. doi:10.1371/journal.pone.0021919
6. Wagner B, Blower S (2010) Costs of eliminating HIV in South Africa have been underestimated. Lancet 376: 953.
7. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
8. Wilson DP (2012) HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med 9: e1001231. doi:10.1371/journal.pmed.1001231
9. Delva W, Eaton JW, Meng F, Fraser C, White RG, et al. (2012) HIV treatment as prevention: optimising the impact of expanded HIV treatment programmes. PLoS Med 9: e1001258. doi:10.1371/journal.pmed.1001258
10. Bärnighausen T, Salomon JA, Sangrujee N (2012) HIV treatment as prevention: issues in economic evaluation. PLoS Med 9: e1001263. doi:10.1371/journal.pmed.1001263
11. Youle M, Trueman P, Simpson K (1999) Health economics in HIV disease. A review of the European literature. Phamacoeconomics 15 (Suppl 1): 1–12.
12. Beck EJ, Miners AH, Yolley K (2001) The cost of HIV treatment and care: a global review. Phamacoeconomics 19: 13–39.
13. Beck EJ, Harling G, Gerbase S, DeLay P (2010) The cost of treatment and care for people living with HIV infection: implications of published studies, 1999–2008. Curr Opin HIV AIDS 5: 215–224.
14. Bollinger L, Adesina A (2011) Review of available ART costs, major cost drivers and potential efficiency gains [presentation]. UNAIDS Economics Reference Group; 7–8 March 2011; Geneva, Switzerland.
15. Galárraga O, Wirtz VJ, Figuero-Lara A, Santa-Ana-Tellez Y, Coulibaly I (2011) Unit costs for delivery of antiretroviral treatment and prevention of mother-to-child transmission of HIV. A systematic review for low- and middle-income countries. Pharmacoeconomics 29: 579–599.
16. Oddone E, Cowper P, Hamilton J, Matchar DB, Hartigan P, et al. (1993) Cost effectiveness analysis of early zidovudine treatment of HIV infected patients. BMJ 307: 1322–1325.
17. Schulman K, Lynne L, Glick H, Eisenberg J (1991) Cost-effectiveness of low-dose zidovudine therapy for asymptomatic patients with human immunodeficiency virus (HIV) infection. Ann Intern Med 114: 798–801.
18. Davies D, Carne C, Camilleri-Ferrante C (1999) Combined antiviral treatment in HIV infection. Is it value for money? Public Health 113: 315–317.
19. Chancellor JV, Hill AM, Sabin CA, Simpson KN, Youle M (1997) Modelling the cost effectiveness of lamivudine/zidovudine combination therapy in HIV infection. Pharmacoeconomics 12: 54–66.
20. Mauskopf J, Lacey L, Kempel A, Simpson K (1998) The cost-effectiveness of treatment with lamivudine and zidovudine compared with zidovudine alone: a comparison of Markov model and trial data estimates. Am J Manag Care 4: 1004–1012.
21. Simpson K, Hatziandreu EJ, Andersson F, Shakespeare A, Oleksy I, et al. (1994) Cost effectiveness of antiviral treatment with zalcitabine plus zidovudine for AIDS patients with CD4+ counts less than 300/microliters in 5 European countries. Pharmacoeconomics 6: 553–562.
22. Biddle AK, Simpson KN (2000) Modeling the use of triple combination therapy in five countries: nevirapine, zidovudine, and didanosine. Value Health 3: 186–201.
23. Sendi PP, Bucher H, Harr T, Craig BA, Schwietert M (1999) Cost effectiveness of highly active antiretroviral therapy in HIV-infected patients. AIDS 13: 1115–1122.
24. Cook J, Dasbach E, Coplan P, Markson L, Yin D, et al. (1999) Modeling the long-term outcomes and costs of HIV antiretroviral therapy using HIV RNA levels: applications to a clinical trial. AIDS Res Hum Retroviruses 15: 499–508.
25. Trueman P, Youle M, Sabin CA, Miners AH, Beck EJ (2000) The cost-effectiveness of triple nucleoside analogue therapy antiretroviral regimens in the treatment of HIV in the United Kingdom. HIV Clin Trials 1: 27–35.
26. Miners A, Sabin C, Trueman P, Youle M, Mocroft A, et al. (2001) Assessing the cost-effectiveness of highly active antiretroviral therapy for adults with HIV in England. HIV Med 2: 52–58.
27. Kahn JG, Haile B, Kates J, Chang S (2001) Health and federal budgetary effects of increasing access to antiretroviral medications for HIV by expanding Medicaid. Am J Public Health 91:1464–1473.
28. Risebrough N, Oh P, Rachlis A, McMurchy D, Bast M, et al. (1999) Economic evaluation of triple ART with indinavir or abacavir and ZDV+3TC compared to dual therapy ZDV+3TC [abstract]. 6th Conference on Retroviruses and Opportunistic Infections; 31 Jan–4 Feb 1999; Chicago, Illinois, US.
29. Caro J, O'Brien J, Miglaccio-Walle K, Raggio G (2001) Economic analysis of initial HIV treatment: efavirenz- versus indinavir-containing triple therapy. Pharmacoeconomics 19: 95–104.
30. Schackman BR, Freedberg KA, Weinstein MC, Sax PE, Losina E, et al. (2002) Cost-effectiveness implications of the timing of antiretroviral therapy in HIV-infected adults. Arch Intern Med 162: 2478–2486.
31. Schackman BR, Goldie SJ, Weinstein MC, Losina E, Zhang H, et al. (2001) Cost-effectiveness of earlier initiation of antiretroviral therapy for uninsured HIV-infected adults. Am J Public Health 91: 1456–1463.
32. Yazdanpanah Y, Goldie S, Losina E, Weinstein MC, Lebrun T, et al. (2002) Lifetime cost of HIV care in France during the era of highly active antiretroviral therapy. Antivir Ther 7: 257–266.
33. Freedberg KA, Losina E, Weinstein MC, Paltiel D, Cohen C, et al. (2001) The cost effectiveness of combination antiretroviral therapy for HIV disease. N Engl J Med 344: 824–831.
34. Mauskopf JA, Tolson JM, Simpson KN, Pham SV, Albright J (2000) Impact of zidovudine-based triple combination therapy on an AIDS drug assistance program. J Acquir Immune Defic Syndr 23: 302–313.
35. Moore RD, Bartlett J (1996) Combination antiretroviral therapy in HIV infection: an economic perspective. Pharmacoeconomics 10: 109–113.
36. Simpson KN, Luo MP, Chumney E, Sun E, Brun S, et al. (2004) Cost-effectiveness of lopinavir/ritonavir versus nelfinavir as the first-line highly active antiretroviral therapy regimen for HIV infection. HIV Clin Trials 5: 294–304.
37. Munakata J, Sanders G, Owens D, Bayoumi A (2003) Cost effectiveness of enfuvirtide in the treatment of drug-resistant HIV infection. Med Decis Making 23: 569.
38. Snedecor S, Hartzema A, Schiller K (2005) Cost effectiveness of HIV treatment innovations of greater efficacy than highly active antiretroviral therapy (HAART). Value Health 8: 244.
39. Sax P, Losina E, Weinstein M, Paltiel A, Goldie S, et al. (2005) Cost-effectiveness of enfuvirtide in treatment-experienced patients with advanced HIV disease. J Acquir Immune Defic Syndr 39: 69–77.
40. Long E, Brandeau M, Galvin C, Vinichenko T, Tole S, et al. (2006) Effectiveness and cost-effectiveness of strategies to expand antiretroviral therapy in St. Petersburg, Russia. AIDS 20: 2207–2215.
41. Goldie S, Yazdanpanah Y, Losina E, Weinstein M, Anglaret X, et al. (2006) Cost-effectiveness of HIV treatment in resource-poor settings—the case of Cote d'Ivoire. N Engl J Med 355: 1141–1153.

Acknowledgments

We thank Christine Michaels for help with the review of economic analyses of ART provision.

Author Contributions

Conceived and designed the experiments: GMR MO. Analyzed the data: GMR MO. Wrote the first draft of the manuscript: GMR MO. Contributed to the writing of the manuscript: GMR MO. ICMJE criteria for authorship read and met: GMR MO. Agree with manuscript results and conclusions: GMR MO.

42. Paton N, Chapman C, Sangeetha S, Mandalia S, Bellamy R, et al. (2006) Cost and cost-effectiveness of antiretroviral therapy for HIV infection in Singapore. Int J STD AIDS 17: 699–705.
43. Cleary S, McIntyre D, Boulle A (2006) The cost-effectiveness of antiretroviral treatment. Cost Eff Resour Alloc 4: 20.
44. Over M, Revenga A, Msasaki E, Peerapatanapokin W, Gold J, et al. (2007) The economics of effective AIDS treatment in Thailand. AIDS 21(Suppl 4): S105–S116.
45. Walensky RP, Wood R, Ciaranello AL, Paltiel AD, Lorenzana SB, et al. (2010) Scaling up the 2010 World Health Organization HIV treatment guidelines in resource-limited settings: a model-based analysis. PLoS Med 7: e1000382. doi:10.1371/journal.pmed.1000382
46. Bendavid E, Grant P, Talbot A, Owens DK, Zolopa A (2011) Cost-effectiveness of antiretroviral regimens in the World Health Organization's treatment guidelines: a South African analysis. AIDS 25: 211–220.
47. Ciaranello AL, Lockman S, Freedberg KA, Hughes M, Chu J, et al. (2011) First-line antiretroviral therapy after single-dose nevirapine exposure in South Africa: a cost-effectiveness analysis of the OCTANE trial. AIDS 25: 479–492.
48. Bachmann MO (2006) Effectiveness and cost effectiveness of early and late prevention of HIV/AIDS progression with antiretrovirals or antibiotics in Southern African adults. AIDS Care 18: 109–120.
49. Bonnel R (2000) Costs of scaling HIV program activities to a national level in sub-Saharan Africa: methods and estimates. Washington (District of Columbia): World Bank.
50. Kumaranayake L, Conteh L, Kurowski C, Watts C (2001) Preliminary estimates of the cost of expanding TB, malaria and HIV/AIDS activities for sub-Saharan Africa. Geneva: Working Group 5, WHO Commission on Macroeconomics and Health.
51. Creese A, Floyd K, Alban A, Guinness L (2002) Cost-effectiveness of HIV/AIDS interventions in Africa: a systematic review of the evidence. Lancet 359: 1635–1642.
52. Hogan DR, Baltussen R, Hayashi C, Lauer JA, Salomon JA (2005) Achieving the millennium development goals for health: cost effectiveness analysis of strategies to combat HIV/AIDS in developing countries. BMJ 331: 1431–1437. doi:10.1136/bmj.38643.368692.68
53. Floyd K, Gilks C (1997) Cost and financing aspects of providing antiretroviral therapy. In: Van Praag E, Fernyak S, Martin Katz A, editors. The implications of antiretroviral treatments. Informal consultation. World Health Organization, Office of HIV/AIDS and Sexually Transmitted Diseases. Geneva: World Health Organization.
54. Hogg R, Weber A, Craib K, Aslam A, O'Shaughnessy M, et al. (1998) One world, one hope: the cost of providing antiretroviral therapy to all nations. AIDS 12: 2203–2209.
55. Attaran A, Sachs J (2001) Defining and refining international donor support for combating the AIDS epidemic. Lancet 357: 57–61
56. Schwartländer B, Stover J, Walker N, Bollinger L, McGreevey W, et al. (2001) Resource needs for HIV/AIDS. Science 292: 2434–2436.
57. Gutierrez J, Johns B, Adam T, Bertozzi SM, Edejer TT (2004) Achieving the WHO/UNAIDS antiretroviral treatment 3 by 5 goal: what will it cost? Lancet 364: 63–64.
58. Stover J, Korenromp E, Blakley M, Komatsu R, Viisainen K (2011) Long-term costs and health impact of continued global fund support for antiretroviral therapy. PLoS ONE 6: e21048. doi:10.1371/journal.pone.0021048
59. Resch S, Korenromp E, Stover J, Blakley M, Krubiner C, et al. (2011) Economic returns to investment in AIDS treatment in low and middle income countries. PLoS ONE 6: e25310. doi:10.1371/journal.pone.0025310
60. Schwartländer B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041.
61. Long EF, Brandeau ML, Owens DK (2010) The cost-effectiveness and population outcomes of expanded HIV screening and antiretroviral treatment in the United States. Ann Intern Med 153: 778–789.
62. Over M, Heywood P, Gold J, Gupta I, Hira S, et al. (2004) HIV/AIDS treatment and prevention in India: modeling the costs and consequences. Washington (District of Columbia): The International Bank for Reconstruction and Development/The World Bank.
63. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, et al. (2007) Cost effectiveness of alternative strategies for initiating and monitoring highly active antiretroviral therapy in the developing world. J Acquir Immune Defic Syndr 46: 91–100.
64. Long L, Fox M, Sanne I, Rosen S (2010) The high cost of second-line antiretroviral therapy for HIV/AIDS in South Africa. AIDS 24: 915–919.
65. Brennan A, Meyer-Rath G, Fox M, Modisenyane T, Martison N (2011) Rates and cost of hospitalization before and after initiation of antiretroviral therapy in South Africa [abstract]. International Workshop on HIV Observational Databases; 24–26 March 2011; Prague, Czech Republic.
66. Krentz HB, Auld MC, Gill MJ (2004) The high cost of medical care for patients who present late (CD4 <200 cells/microL) with HIV infection. HIV Med 5: 93–98.
67. Nykamp D, Barnett CW, Lago M, Parham DL, Fernandez ES (1997) Cost of medication therapy in ambulatory HIV-infected patients. Ann Pharmacother 31: 303–307.
68. Leisegang R, Cleary S, Hislop M, Davidse A, Regensberg L, et al. (2009) Early and late direct costs in a Southern African antiretroviral treatment programme: a retrospective cohort analysis. PLoS Med 6: e1000189. doi:10.1371/journal.pmed.1000189
69. Perez-Casas C, Herranz E, Ford N (2001) Pricing of drugs and donations: options for sustainable equity pricing. Trop Med Int Health 6: 960–964.
70. World Trade Organization (2001) Declaration on the TRIPS agreement and public health. WT/MIN(01)/DEC/2.
71. Médecins Sans Frontières Campaign for Access to Essential Medicines (2008) Untangling the web of antiretroviral price reductions, 11th edition. Geneva: Médecins Sans Frontières.
72. Joint United Nations Programme on HIV/AIDS (2011) Toward an improved framework for HIV investments. Geneva: Joint United Nations Programme on HIV/AIDS.
73. Marseille E, Dandona L, Marshall N, Gaist P, Bautista-Arradondo S, et al. (2007) HIV prevention cost and program scale: data from the PANCEA project in five low and middle-income countries. BMC Health Serv Res 7: 108.
74. Guinness L, Kumaranayake L, Rajaraman B, Sankaranarayanan G, Vannela G, et al. (2005) Does scale matter? The costs of HIV-prevention interventions for commercial sex workers in India. Bull World Health Organ 83: 747–755.
75. Guinness L, Kumaranayake L, Hanson K (2007) A cost function for HIV prevention services: is there a 'u'-shape? Cost Eff Resour Alloc 5: 13.
76. Johns B, Baltussen R (2004) Accounting for the cost of scaling-up health interventions. Health Econ 13: 1117–1124.
77. Menzies NA, Berruti AA, Berzon R, Filler S, Ferris R (2011) The cost of providing comprehensive HIV treatment in PEPFAR-supported programs. AIDS 25: 1753–1760.
78. Rosen S, Long L, Sanne I (2008) The outcomes and outpatient costs of different models of antiretroviral treatment delivery in South Africa. Trop Med Int Health 13: 1005–1015.
79. Kabugo C, Bahendeka S, Mwebaze R, Malamba S, Katuntu D, et al. (2005) Long-term experience providing antiretroviral drugs in a fee-for-service HIV clinic in Uganda. Evidence of extended virologic and CD4+ cell count responses. J Acquir Immune Defic Syndr 38: 578–583.
80. Leonard KL, Masatu MC (2010) Professionalism and the know-do gap: exploring intrinsic motivation among health workers in Tanzania. Health Econ 19: 1461–1477.
81. Das J, Hammer J, Leonard K (2008) The quality of medical advice in low-income countries. J Econ Perspect 22: 93–114.
82. Republic of South Africa (2010) Country progress report on the declaration of commitment on HIV/AIDS. 2010 Report. Pretoria: Republic of South Africa.
83. Zipf GK (1935) The psychobiology of language. Boston: Houghton-Mifflin.

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Issues in Economic Evaluation

Till Bärnighausen[1,2]*, Joshua A. Salomon[1], Nalinee Sangrujee[3]

1 Department of Global Health and Population, Harvard School of Public Health, Boston, Massachusetts, United States of America, **2** Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Mtubatuba, KwaZulu-Natal, South Africa, **3** Division of Global HIV/AIDS, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America

Abstract: Meyer-Rath and Over assert in another article in the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections", that economic evaluations of antiretroviral therapy (ART) in currently existing programs and in HIV treatment as prevention (TasP) programs should use cost functions that capture cost dependence on a number of factors, such as scale and scope of delivery, health states, ART regimens, health workers' experience, patients' time on treatment, and the distribution of delivery across public and private sectors. We argue that for particular evaluation purposes (e.g., to establish the social value of TasP) and from particular perspectives (e.g., national health policy makers) less detailed cost functions may be sufficient. We then extend the discussion of economic evaluation of TasP, describing why ART outcomes and costs assessed in currently existing programs are unlikely to be generalizable to TasP programs for several fundamental reasons. First, to achieve frequent, widespread HIV testing and high uptake of ART immediately following an HIV diagnosis, TasP programs will require components that are not present in current ART programs and whose costs are not included in current estimates. Second, the early initiation of ART under TasP will change not only patients' disease courses and treatment experiences—which can affect behaviors that determine clinical treatment success, such as ART adherence and retention—but also quality of life and economic outcomes for HIV-infected individuals. Third, the preventive effects of TasP are likely to alter the composition of the HIV-infected population over time, changing its biological and behavioral characteristics and leading to different costs and outcomes for ART.

More versus Less Detailed Cost Functions

The results from the HTPN 052 trial reported in August 2011 demonstrated under the controlled conditions of a well-conducted clinical trial that early antiretroviral therapy (ART) can be highly effective in preventing transmission of HIV in stable heterosexual HIV-discordant couples [1]. Several experimental studies are currently underway or planned to investigate the effectiveness of HIV treatment as prevention (TasP) in general populations, including in HIV hyperendemic communities in sub-Saharan Africa [2,3]. A few mathematical modeling studies have predicted the cost-effectiveness of TasP, using cost estimates derived from currently existing ART programs [4–7]. Meyer-Rath and Over review prior studies of ART costs, and discuss the cost assumptions used in economic evaluations of HIV treatment [8]. They find that economic evaluations of TasP have tended toward a simplified accounting for variation in ART costs across patients and settings, focusing on a limited set of factors such as regimen or disease stage. Meyer-Rath and Over argue that future economic evaluation should account for a range of other factors that may be significant determinants of ART costs, including scale and scope of delivery, health states, ART regimens, health workers' experience, patients' time on treatment, and the distribution of delivery across public and private sectors.

In making this argument, Meyer-Rath and Over distinguish between two categories of ART cost functions [8]: "cost accounting identities," which generate estimates of total costs based on mathematical representations of the production process, and "flexible cost functions," which generate estimates of total costs based on empirically derived relationships between costs and other factors, while treating the details of the production process as a "black box" (Text S1 of [8]). Meyer-Rath and Over find that that "[m]ost existing [ART] cost projections assume a single constant unit cost per patient-year, or per patient-year on a certain regimen," while a few have allowed for variation of costs by disease stage but not by other factors [8]. Concerns with the level of detail in modeling the costs of TasP derive in part from the past focus on predictive, or *ex ante*, economic evaluations, which rely heavily on mathematical or statistical models to extrapolate from limited empirical observations (as opposed to *ex post* evaluations, which use direct observation of actual costs and benefits) [9].

It may indeed be ideal to capture the dependence of costs on many factors in economic evaluation of TasP—a task that could theoretically be achieved either by improving our understanding of the production process or through empirical examination of relationships between costs and other factors. However, the necessary data on the ART production process or on the relationship between ART costs and factors such as the scope of delivery or patients' time on treatment are currently largely lacking

Citation: Bärnighausen T, Salomon JA, Sangrujee N (2012) HIV Treatment as Prevention: Issues in Economic Evaluation. PLoS Med 9(7): e1001263. doi:10.1371/journal.pmed.1001263

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: TB received funding support from National Institute of Child Health and Human Development (1R01-HD058482-01), the National Institute of Mental Health (1R01MG083539-01), World Bank, and the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Joshua Salomon is a member of the *PLoS Medicine* Editorial Board. We declare that we have no other conflicts of interest.

Abbreviations: ART, antiretroviral therapy; TasP, HIV treatment as prevention

* E-mail: tbaernig@hsph.harvard.edu

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

and may not become widely available for most settings in the near future, despite ongoing studies that will generate such data for a few settings. The absence of empirical data raises the question whether economic evaluation of TasP can be "good enough" without accounting for the dependence of ART costs on many of the factors that Meyer-Rath and Over argue convincingly could be determinants of ART costs.

The answer to this question will depend on both evaluation purpose and perspective. If the purpose is to decide whether or not to implement TasP, less detailed cost functions may be sufficient, because the result will be a yes/no answer indicating whether TasP produces a net benefit to society, or falls below some predetermined cost-effectiveness threshold. Such a result may be relatively robust to imprecision in the specification and estimation of costs. If, on the other hand, the purpose of the evaluation is to establish the most efficient approach to deliver TasP, given that it has been decided that it should be implemented, it will be crucial for the analysis to capture cost variations based on factors such as health worker–to–patient ratio, size and type of health care facility, and the level of integration of TasP programs into the general health care system.

The example Meyer-Rath and Over calculate in their article is a case in point. Based on theoretical considerations of economies of scale and empirical observation of scale effects in most industries, including in the delivery of HIV prevention services [10–12], they argue that it is unlikely that average costs would remain constant across scale of ART delivery. To demonstrate the potential impact of scale effects on costs, they adjust the estimates for implementing TasP in South Africa produced by Granich et al. [13] "for scale and a plausible pattern of distribution of patients into clinics" [8]. The result of this adjustment is an increase in total accumulated cost over 40 years from US$75 billion to US$106 billion. While this difference in cost estimates is large, given the dramatic effect of TasP estimated by Granich et al. [13], it may not alter the overall conclusion that TasP is a socially worthwhile intervention. In particular, if TasP could indeed eliminate HIV incidence, as Granich et al. [13] assert based on their modeling results, the economic case for TasP would likely be robust to large increases in the cost estimates. At the same time, cost increases of the magnitude calculated by Meyer-Rath and Over would be extremely important for the practical exercises of financial planning and budgeting for TasP.

Whether less detailed cost functions will suffice in a given situation is also affected by the evaluation perspective. For instance, from the perspective of national health policy makers, economic evaluation may not need to account for the relationship between ART costs and the sector of delivery because policy makers may not be concerned about patients who utilize ART in the private sector. Conversely, from the perspective of for-profit companies providing ART in workplace HIV treatment programs, only the private sector costs will be relevant. In such cases, cost functions that account for differences in public versus private sector costs are not required.

Meyer-Rath and Over start the important discussion of which factors to include in cost functions in economic evaluations of TasP. We argue that for particular evaluation purposes (e.g., to establish the social value of TasP) and from particular perspectives (e.g., national health policy makers) "undetailed" cost functions, which do not capture cost dependence on many factors, may be sufficient.

ART and TasP Programs: Important Considerations for Economic Evaluation

A more fundamental issue in the economic evaluation of TasP is the fact that TasP programs will differ from existing ART programs in several important respects, which may substantially impact the costs, outcomes, and cost-effectiveness of TasP (for an overview, see Table 1).

Program Components and Costs

As Meyer-Rath and Over observe, estimates of ART costs in models predicting the cost-effectiveness of TasP are usually based on the mix of interventions in currently existing ART programs. However, TasP strategies will likely require interventions that are not usually present in existing ART programs [14]. The costs of these interventions need to be accounted for in economic evaluations of TasP. As data on CD4 count at ART initiation demonstrate [15,16], most HIV-infected people in sub-Saharan Africa do not start treatment until they have reached advanced stages of HIV disease (despite a moderate but steady increase in median CD4 count at initiation over the past years [16]). The late ART initiation occurs despite national treatment guidelines in sub-Saharan Africa stipulating that all HIV-positive individuals be

Table 1. Overview of differences between existing ART programs and TasP.

Difference	Explanation
Program components	To achieve widespread and frequent HIV testing and high uptake of ART immediately following an HIV diagnosis, TasP strategies will likely require interventions that are not present in current ART programs.
Disease experiences and treatment outcomes	Disease experiences and treatment-relevant behaviors: patients who initiate ART early are less likely to experience recovery from the symptoms of the later stages of HIV disease. Lack of such experience may affect behaviors that are crucial for treatment outcomes, such as ART retention and adherence.
	Quality of life: early initiation may reduce quality of life (because it increases the duration of drug side effects and transforms people into patients several years earlier than under current ART guidelines) or improve quality of life (because it decreases the severity of drug side effects and avoids the psychologically distressing situation of having to wait for one's health status to deteriorate before being allowed to start ART).
	Economic productivity: early initiation may reduce the lifetime economic productivity of HIV-infected individuals (because it increases the total portion of lifetime spent utilizing ART) or improve productivity (because it avoids the negative economic consequences of deteriorating health preceding late ART initiation).
Patient population	In the long run, successful TasP could lead to changes in the composition of the people requiring ART, because the preventive effects of TasP may benefit some population subgroups at risk of HIV acquisition, but not others.

doi:10.1371/journal.pmed.1001263.t001

referred to CD4 count testing and clinical disease staging after a positive HIV test result [17].

To succeed in reducing HIV incidence in the general population, TasP programs require that very high proportions of all HIV-infected people in a community receive ART. To this end, TasP programs must ensure that most community members who have not been diagnosed with HIV frequently test for HIV and, if found to be HIV-infected, initiate treatment soon after diagnosis. The costs of interventions to achieve frequent, widespread HIV testing (e.g., through community mobilization and home-based testing) and high uptake of ART among people newly diagnosed with HIV (e.g., through counseling, improved transport, and financial incentives) must be included on the cost side of the economic evaluation of TasP [18].

In addition, some structures and processes that are in place in existing ART programs may need to be enhanced to achieve good TasP outcomes, e.g., community-based treatment supporters or mobile-phone short messages to ensure good ART retention and adherence [19]. The expenditures for such enhancements need to be accounted for in the economic evaluation of TasP. Of course, the specific components required for successful TasP will depend on the particular TasP intervention strategy—universal population-wide HIV testing and treatment will use different approaches, and incur different costs, than TasP strategies targeted at people at high risk of HIV transmission, such as HIV-infected individuals in HIV-discordant couples [2,13,20]. It will thus be crucial that TasP implementations both in trials and in routine settings are accompanied by rigorous empirical studies that measure the expenditures for all TasP components.

Disease Experiences and Treatment Outcomes

A second issue in predicting the cost-effectiveness of TasP based on costs and outcomes observed in current ART programs is that TasP patients are initiated earlier on ART than patients in existing programs. While patients in sub-Saharan Africa currently initiate ART at a median CD4 cell count below 140 cells/μl [15] (which is substantially lower than the typical CD4 count eligibility thresholds of 200 cells/μl or 350 cells/μl), patients in successful TasP programs would initiate ART soon after first diagnosis of HIV infection (which ideally would occur soon after HIV infection). As a result of earlier ART initiation, patients' disease experiences and treatment outcomes are likely to be significantly altered.

Disease experiences and treatment-relevant behaviors. Patients who initiate ART early are unlikely to experience the symptoms of the later stages of HIV disease, as well as the subsequent recovery on treatment, which many patients enrolled in currently existing ART programs have experienced. The recovery on ART from weight loss, physical weakness, and the opportunistic infections of late-stage HIV disease may convince patients in current treatment programs that ART is indeed effective and, as a result, improve their long-term ART retention and adherence. Patients in successful TasP programs, on the other hand, will usually lack such experiences and may consequently be less motivated to adhere well to their clinical appointments and drug regimens. Rates of resistance development, mortality, and morbidity may thus be higher in TasP programs than in existing ART programs.

Quality of life. TasP is also likely to affect quality of life, but the direction and magnitude of the net effect over a patient's life course is unknown. TasP patients who initiate ART early will experience drug side effects of ART for a longer total duration than patients in currently existing programs. At the same time, side effects and toxicities may be less frequent or less severe in patients who initiate ART early [21,22], and TasP patients who adhere well to their treatment regimens may be able to completely avoid some of the symptoms of more advanced stages of HIV disease that can substantially reduce quality of life [23]. It is further plausible that early ART initiation improves quality of life because it avoids the psychologically distressing situation of having to wait for one's health status to deteriorate before being allowed to start ART. Conversely, it is also plausible that early initiation reduces quality of life, because it transforms people with no obvious symptoms into patients, and burdens them with the responsibilities of chronic disease treatment, such as regular clinic visits and pill-taking, several years earlier than under current ART guidelines.

Importantly, the net effect of TasP on quality of life will depend on the counterfactual to which it is compared. Since economic evaluation is intended to inform the decision whether to implement TasP against the background of already existing policies, the best counterfactual will be ART provided at the current ART eligibility threshold in a country. In particular settings, such as the US, where individuals are currently already eligible at the highest CD4 count for which there is clear evidence of health benefit to the HIV-infected patient [24], TasP will be equivalent to initiating on ART a group of people who will not yet derive benefits for their own health from the treatment. In these situations, limiting the health effects of TasP will be limited to reduction of HIV transmission to others and quality of life changes. Empirical evaluations of TasP should thus always include quality of life assessments.

Economic productivity. In current ART programs, ART patients' economic situation commonly improves with time on treatment [25–27], and recent evidence from a population-based study shows that ART can lead to nearly full employment recovery among HIV patients in rural Southern Africa [28]. In the context of TasP, patients who adhere well to their treatment regimens will be unlikely to experience negative effects of HIV on economic productivity because they initiate ART many years before they would have suffered from significant HIV disease, had they not received ART. On the other hand, TasP patients will start incurring the time losses and transport costs of ART utilization several years earlier in their disease course than patients enrolled in currently existing ART programs. For economic evaluations that take the perspective of the society as a whole, which include patients' private expenditure and economic productivity in addition to the costs incurred by the public health care sector, ART effects on patient income must be incorporated in the analysis. As the direction and magnitude of TasP effects on economic outcomes over a patient's life course are currently unknown, these effects need to be established in empirical studies of TasP.

Changes in the HIV-Infected Population

Above, we have argued that the same people would behave differently in TasP programs than they currently do in existing ART programs. But, over time, TasP will also change who the people living with HIV are. One important reason for this change arises because people who newly acquire HIV despite successful implementation of TasP in their communities are likely to differ in their biology or behavior from the people who have acquired HIV in the past and are currently receiving treatment. If TasP is indeed effective in averting onward transmission of HIV, the people who acquire HIV in the presence of TasP may have particularly weak immune-system functioning or engage in exceptionally high-risk sexual behavior with HIV-infected people who do not participate in TasP. Thus, under successful TasP strategies the population of HIV-infected individuals will not only be smaller (as the TasP

effects reducing transmission will over time outweigh the effects on survival [29,30]), but it will also possess different average characteristics, e.g., regarding biological susceptibility to HIV infection or sexual risk-taking. As average ART patient characteristics change following the introduction of TasP, it is likely that ART costs and outcomes will change as well, because some of the characteristics that determine the extent to which a person is protected from HIV acquisition by TasP will also affect ART success (such as immunological functioning or adherence behavior).

Conclusion

Over the coming years, more detailed cost data are likely to become available, which will allow incorporating the cost determinants that Meyer-Rath and Over identify into predictive models of the economic value of TasP. While these data are not yet available, it is important to keep in mind that detail in representing the dependence of ART costs on a range of factors in economic evaluation is likely to matter far less for establishing whether TasP is beneficial for a population than for determining which models of TasP delivery allocate scarce resources optimally. For the former purpose, relatively simple cost functions may be sufficient.

Several more fundamental issues deserve consideration in setting up economic evaluations of TasP. First, to achieve frequent, widespread HIV testing and high uptake of ART immediately following an HIV diagnosis, TasP programs will likely require components that are not present in current ART programs and whose costs are thus not incorporated in current cost estimates (such as community mobilization and frequent HIV testing of all community members who have not been diagnosed with HIV). Second, the early initiation of ART under TasP will change not only patients' disease courses and treatment experiences (which can affect behaviors that determine clinical treatment success, such as ART adherence and retention), but also the quality of life and economic productivity of HIV-infected populations—changes in outcomes that need to be accounted for in economic evaluation. And, third, the preventive effects of TasP are likely to alter the composition of the HIV-infected population in the long run, changing its biological and behavioral characteristics and leading to ART costs and outcomes that are different from those observed for current ART patients.

Key Points

- Using antiretroviral treatment (ART) cost functions that capture a range of relationships between costs and other factors, as recommended by Meyer-Rath and Over in another article in this collection [8], will likely be important for economic evaluations aiming to determine which delivery model of HIV treatment as prevention (TasP) will allocate resources optimally, but it may not be necessary to establish whether TasP will increase social welfare.
- There are several problems inherent in predicting the cost-effectiveness of TasP based on outcomes and costs observed in currently existing ART programs, because TasP will (i) require components that are not present in existing programs, (ii) affect the disease courses and treatment experiences, quality of life, and economic productivity of people living with HIV, and (iii) change the composition of the HIV-infected population.
- To improve our capacity to estimate the social value of TasP in economic evaluation, empirical studies need to comprehensively assess the costs of all TasP components and measure not only the TasP effects on HIV incidence and mortality but also the impacts on quality of life and economic productivity.

Hence, while it is useful to predict the economic value of TasP using data from existing ART programs, such forecasts can only serve as an initial guide for health policy. Future studies accompanying ongoing and planned TasP trials and implementations need to comprehensively assess the costs of all TasP components, as well as TasP effects on a range of outcomes beyond HIV incidence and mortality, including quality of life and economic productivity.

Author Contributions

Analyzed the data: TB JAS. Contributed reagents/materials/analysis tools: TB JAS NS. Wrote the first draft of the manuscript: TB. Contributed to the writing of the manuscript: JAS NS. ICMJE criteria for authorship read and met: TB JAS NS. Agree with manuscript results and conclusions: TB JAS NS.

References

1. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
2. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469.
3. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250
4. Hontelez JA, de Vlas SJ, Tanser F, Bakker R, Bärnighausen T, et al. (2011) The impact of the new WHO antiretroviral treatment guidelines on HIV epidemic dynamics and cost in South Africa. PLoS ONE 6: e21919. doi:10.1371/journal.pone.0021919
5. Long EF, Brandeau ML, Owens DK (2010) The cost-effectiveness and population outcomes of expanded HIV screening and antiretroviral treatment in the United States. Ann Intern Med 153: 778–789.
6. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, et al. (2007) Cost-effectiveness of alternative strategies for initiating and monitoring highly active antiretroviral therapy in the developing world. J Acquir Immune Defic Syndr 46: 91–100.
7. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
8. Meyer-Rath G, Over M (2012) HIV treatment as prevention: modelling the cost of antiretroviral treatment—state of the art and future directions. PLoS Med 9: e1001247. doi:10.1371/journal.pmed.1001247
9. Bärnighausen T, Kyle M, Salomon JA, Waning B (2011) Assessing the population health impact of market interventions to improve access to antiretroviral treatment. Health Policy Plan. E-pub ahead of print. doi:10.1093/heapol/czr058
10. Guinness L, Kumaranayake L, Hanson K (2007) A cost function for HIV prevention services: is there a 'u' - shape? Cost Eff Resour Alloc 5: 13.
11. Guinness L, Kumaranayake L, Rajaraman B, Sankaranarayanan G, Vannela G, et al. (2005) Does scale matter? The costs of HIV-prevention interventions for commercial sex workers in India. Bull World Health Organ 83: 747–755.
12. Marseille E, Dandona L, Marshall N, Gaist P, Bautista-Arredondo S, et al. (2007) HIV prevention costs and program scale: data from the PANCEA project in five low and middle-income countries. BMC Health Serv Res 7: 108.
13. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
14. Hallett TB, Gregson S, Dube S, Mapfeka ES, Mugurungi O, et al. (2011) Estimating the resources required in the roll-out of universal access to antiretroviral treatment in Zimbabwe. Sex Transm Infect 87: 621–628.
15. Nash D, Wu Y, Elul B, Hoos D, El Sadr W (2011) Program-level and contextual-level determinants of low-median CD4+ cell count in cohorts of persons initiating ART in eight sub-Saharan African countries. AIDS 25: 1523–1533.

16. Lahuerta M, Hoffman S, Elul B, Wu Y, Gorrell Kulkarni S, et al. (2012) Change over time in CD4+ count and disease stage at entry into HIV care—findings from 9 sub-Saharan African countries [abstract]. 19th Conference on Retroviruses and Opportunistic Infections; 5–8 March 2012; Seattle, Washington, US.
17. South African Department of Health (2010) HIV counselling and testing (HCT) policy guidelines. Pretoria: South African Department of Health.
18. Bärnighausen T, Tanser F, Dabis F, Newell M (2012) Interventions to improve the performance of HIV health systems for treatment-as-prevention in sub-Saharan Africa: the experimental evidence. Curr Opin HIV AIDS 11: 942–951.
19. Bärnighausen T, Chaiyachati K, Chimbindi N, Peoples A, Haberer J, et al. (2011) Interventions to increase antiretroviral adherence in sub-Saharan Africa: a systematic review of evaluation studies. Lancet Infect Dis 11: 942–951.
20. Hallett TB, Baeten JM, Heffron R, Barnabas R, de Bruyn G, et al. (2011) Optimal uses of antiretrovirals for prevention in HIV-1 serodiscordant heterosexual couples in South Africa: a modelling study. PLoS Med 8: e1001123. doi:10.1371/journal.pmed.1001123
21. Sterne J, May M, Costagliola D, Egger M, Hogg R, et al. (2006) Early HAART associated with improved outcomes & fewer toxicities [abstract]. 13th Conference on Retroviruses and Opportunistic Infections; 5–8 February 2006; Denver, Colorado, US.
22. Lichtenstein K, Armon C, Buchacz K, Moorman A, Wood K, et al. (2006) Early, uninterrupted ART is associated with improved outcomes and fewer toxicities in the HIV outpatient study (HOPS) [abstract]. 13th Conference on Retroviruses and Opportunistic Infections; 5–8 February 2006; Denver, Colorado, US.
23. Armon C, Lichtenstein K (2011) The associations among coping, nadir CD4+ T-cell count, and non-HIV-related variables with health-related quality of life among an ambulatory HIV-positive patient population. Qual Life Res. E-pub ahead of print. doi:10.1007/s11136-011-0017-2
24. Panel on Antiretroviral Guidelines for Adults and Adolescents (2012) Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Washington (District of Columbia): US Department of Health and Human Services.
25. Thirumurthy H, Graff Zivin J, Goldstein MP (2005) The economic impact of AIDS treatment: labor supply in western Kenya. J Hum Res 43: 511–552.
26. Rosen S, Ketlhapile M, Sanne I, Desilva MB (2008) Differences in normal activities, job performance and symptom prevalence between patients not yet on antiretroviral therapy and patients initiating therapy in South Africa. AIDS 22 (Suppl 1): S131–S139.
27. Larson BA, Fox MP, Rosen S, Bii M, Sigei C, et al. (2008) Early effects of antiretroviral therapy on work performance: preliminary results from a cohort study of Kenyan agricultural workers. AIDS 22: 421–425.
28. Bor J, Tanser F, Newell ML, Bärnighausen T (2012) Nearly full employment recovery among HIV patients on antiretroviral therapy: evidence from a large population cohort. Health Aff (Millwood). In press.
29. Hontelez JA, Lurie MN, Newell ML, Bakker R, Tanser F, et al. (2011) Ageing with HIV in South Africa. AIDS 25: 1665–1667.
30. Hontelez JA, de Vlas SJ, Baltussen R, Bakker R, Tanser F, et al. (2012) The impact of antiretroviral treatment on the age composition of the HIV epidemic in sub-Saharan Africa. AIDS. In press.

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Optimising the Impact of Expanded HIV Treatment Programmes

Wim Delva[1,2]*, Jeffrey W. Eaton[3], Fei Meng[2,4], Christophe Fraser[3], Richard G. White[5], Peter Vickerman[6], Marie-Claude Boily[3], Timothy B. Hallett[3]

1 South African Department of Science and Technology/National Research Foundation Centre for Excellence in Epidemiological Modelling and Analysis, University of Stellenbosch, Stellenbosch, South Africa, **2** International Centre for Reproductive Health, Ghent University, Ghent, Belgium, **3** Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **4** Center for Statistics, Hasselt University, Diepenbeek, Belgium, **5** Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, United Kingdom, **6** Department of Global Health and Development, London School of Hygiene & Tropical Medicine, London, United Kingdom

Abstract: Until now, decisions about how to allocate ART have largely been based on maximising the therapeutic benefit of ART for patients. Since the results of the HPTN 052 study showed efficacy of antiretroviral therapy (ART) in preventing HIV transmission, there has been increased interest in the benefits of ART not only as treatment, but also in prevention. Resources for expanding ART in the short term may be limited, so the question is how to generate the most prevention benefit from realistic potential increases in the availability of ART. Although not a formal systematic review, here we review different ways in which access to ART could be expanded by prioritising access to particular groups based on clinical or behavioural factors. For each group we consider (i) the clinical and epidemiological benefits, (ii) the potential feasibility, acceptability, and equity, and (iii) the affordability and cost-effectiveness of prioritising ART access for that group. In re-evaluating the allocation of ART in light of the new data about ART preventing transmission, the goal should be to create policies that maximise epidemiological and clinical benefit while still being feasible, affordable, acceptable, and equitable.

Introduction

There has been a rapid expansion in access to antiretroviral therapy (ART) over the past decade, especially in the countries with the highest burden of HIV. At the end of 2010, an estimated 6.7 million people were on ART globally, an increase of over 1.4 million from the previous year, but around 7.5 million people are still in need of treatment based on current World Health Organization (WHO) guidelines [1]. Until now, decisions around how to allocate ART have been based on maximising the therapeutic benefit of ART for patients, within the constraints of limited financial and health care system resources [2]. This has led to ART access being prioritised for those with the lowest CD4 cell counts (and patients with active tuberculosis [TB]) [3].

The HPTN 052 study [4] demonstrated that earlier ART initiation can reduce heterosexual HIV transmission [5]. This finding suggests that future expansions of ART access should seek to maximise not only the therapeutic but also the prevention benefits of treatment. Currently, constrained resources and capacity for HIV treatment and prevention [6–8] make it unfeasible to immediately provide ART for all people living with HIV, even if this was the optimal epidemiological and therapeutic strategy and was widely accepted by communities. However, as increasingly high levels of access under current guidelines are achieved in coming years, the recent information about the prevention benefit of ART has inspired renewed discussion about whether and how to incrementally expand access to treatment to subgroups that will differentially benefit from the preventive and therapeutic features of ART, especially in sub-Saharan Africa, where the burden of HIV is greatest.

Candidate priority groups for early treatment are defined by both clinical and behavioural criteria. Potential clinical criteria for providing early treatment include the following: incrementally increasing the CD4 cell count threshold for treatment eligibility, immediate treatment for those with high set-point viral load, immediate treatment for pregnant women, and immediate treatment for those with TB coinfection. Behavioural risk groups that have been proposed for early treatment include HIV-serodiscordant couples, female sex workers (FSWs), men who have sex with men (MSM), and people who inject drugs (PWID). Expanding access to treatment for each of these subgroups is evaluated here according to (i) clinical and epidemiological benefits, (ii) potential feasibility, acceptability, and equity, and (iii) affordability and cost-effectiveness (Box 1).

This article is not a systematic literature review of all clinical, epidemiological, and policy implications of alternative options for

Citation: Delva W, Eaton JW, Meng F, Fraser C, White RG, et al. (2012) HIV Treatment as Prevention: Optimising the Impact of Expanded HIV Treatment Programmes. PLoS Med 9(7): e1001258. doi:10.1371/journal.pmed.1001258

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: WD thanks the Canadian International Development Agency (CIDA), the Flemish Interuniversity Council (VLIR), and the Flemish Scientific Research Fund (FWO) for funding support. FM was funded by the Flemish Scientific Research Fund (FWO). TBH thanks the Bill & Melinda Gates Foundation for funding support. RGW was funded by Medical Research Council (UK) Methodology Research Fellowship (G0802414) and the Consortium to Respond Effectively to the AIDS/TB Epidemic, funded by the Bill & Melinda Gates Foundation (19790.01). PV was funded by the Bill & Melinda Gates Foundation, Medical Research Council, and the World Bank. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: ANC, antenatal care; ART, antiretroviral therapy; FSW, female sex worker; MSM, men who have sex with men; PEPFAR, US President's Emergency Plan for AIDS Relief; PMTCT, prevention of mother-to-child transmission; PWID, people who inject drugs; SPVL, set-point viral load; TB, tuberculosis; WHO, World Health Organization

* E-mail: Wim.Delva@ugent.be

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Box 1. A Breakdown of Questions Related to the Impact, Feasibility, Affordability, and Acceptability of Expanded ART Provision

Epidemiological Impact

- What is the incremental effectiveness of the expanded ART programme for averting new HIV infections, relative to the existing HIV prevention and treatment programme?
- What is the likelihood of behavioural risk substitution that could undermine prevention benefits?

Clinical Impact

- What is the incremental effectiveness of expanding ART for averting HIV-related morbidity and mortality, relative to the existing modes of ART delivery?
- What is the potential impact of expanded access to ART on the acquisition and transmission of drug resistance?

Affordability and Cost-Effectiveness

- What size is the additional priority group?
- What is the expected start-up cost of the expanded ART programme?
- How would the programme costs accumulate over time?
- Would the programme be cost-effective compared to accepted international benchmarks, and relative to alternative HIV prevention methods?

Feasibility

- What infrastructure and human resources does the expanded ART programme require?
- How would the prioritisation for a particular group be operationalized?
- What is the expected adherence and retention in care for the additional priority group?

Acceptability

- Would the expanded ART programme violate principles of health ethics or human rights?
- Would the expanded ART programme be acceptable to the newly eligible priority group, communities, and decision-makers?

expanding HIV treatment programmes. Rather, it represents an organised collection of expert opinions, literature reviews, and multidisciplinary discussions. Following the publication of the HPTN 052 results and the US President's Emergency Plan for AIDS Relief (PEPFAR) Scientific Advisory Board recommendations for PEPFAR HIV treatment programmes [2], experts in the field of HIV epidemiology, mathematical modelling, and HIV policy were convened in an HIV Modelling Consortium (http://www.hivmodelling.org) meeting in November 2011 to discuss the potential impact of expanded HIV treatment in sub-Saharan Africa. Following from this meeting, this review focuses on the biological and behavioural factors that determine the potential impact of various ART prioritisation options on transmission, morbidity, and mortality, as well as the factors affecting feasibility, affordability, acceptability, cost-effectiveness, and health systems interactions. These include the relative size of the priority group, anticipated ease of identification and recruitment of the priority group, treatment uptake, adherence and loss to follow-up, ethical challenges, and technical and human resources required.

Potential Prioritisation Groups for ART Expansion CD4 Cell Count

CD4 Cell Count

As many low- and middle-income countries are moving towards adoption of the WHO guidelines of providing treatment for all HIV-infected individuals with CD4 cell counts less than 350 cells/μl [3], one natural strategy for increasing the prevention benefit of treatment is to further increase the threshold of eligibility to those with CD4 counts less than 500 cells/μl. Current US treatment guidelines recommend initiation of treatment for asymptomatic HIV-infected individuals with CD4 counts between 350 and 500 cells/μl [9], and European guidelines suggest that treatment should be considered at this point [10]. Observational and clinical trial data that link transmission events confirm that heterosexual transmissions occur from asymptomatically infected individuals with CD4 counts between 350 and 500 cells/μl [5,11], and the HPTN 052 study demonstrated a 96% reduction in transmission associated with treatment initiation at a CD4 cell count between 350 and 550 cells/μl compared to delaying treatment until CD4 count was below 250 cells/μl [5]. However, compared to other CD4 strata, individuals with CD4 counts between 350 and 500 cells/μl have the lowest transmission rates [11] (Figure 1), suggesting that expanding treatment to this group without considering other biological or behavioural transmission risk factors may be the least efficient strategy for prioritising treatment for prevention.

The magnitude of the overall long-term additional therapeutic benefit of providing treatment at CD4 count above 350 cells/μl is uncertain. A collaborative analysis of observational data found that deferring treatment initiation from between 351 and 450 cells/μl to between 251 and 350 cells/μl increased the hazard of AIDS or death by 28% [12], and the HPTN 052 trial found that delaying treatment until CD4 count was lower than 250 cells/μl was associated with a 41% increased hazard of adverse clinical outcome [5]. However, the potential benefits of early treatment need to be weighed against the potential toxicities of ART and negative effects on quality of life [9]. Earlier treatment initiation may also be associated with poorer adherence or retention in care [13], which can lead to increased risk of drug-resistant virus. More robust data about the clinical benefit of earlier treatment and patients' retention in care when treatment is initiated earlier are expected from the START trial [14].

Both the cost and epidemiological impact of expanding eligibility for ART to those with CD4 counts up to 500 cells/μl will largely be determined by the number of additional people on treatment. Cross-sectional data from sub-Saharan Africa suggest that between 20% and 25% of HIV-infected people have CD4 counts between 350 and 500 cells/μl [15]. Based on the Joint United Nations Programme on HIV/AIDS estimate of approximately 19.8 million adults infected in sub-Saharan Africa [16], increasing the CD4 threshold would add between 4 and 5 million to the 10 million people currently still in need of treatment.

However, even with a change in the threshold at which patients are considered eligible for treatment, the numbers expected to initiate treatment at high CD4 counts will be low without improvements in frequency of testing and retention in pre-ART care [17]. Surveillance of HIV testing programmes in a township near Cape Town, South Africa, found that amongst individuals accessing voluntary counselling and testing, 66% of those testing HIV-positive already had CD4 cell counts below 350 cells/μl [18].

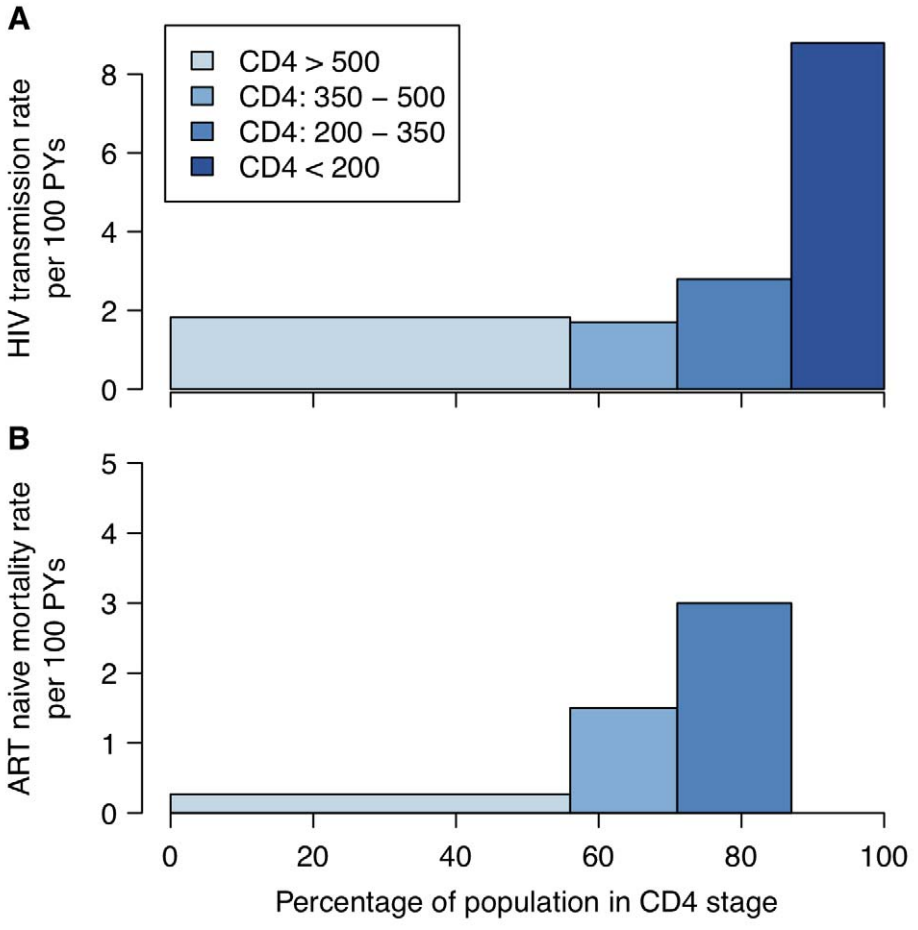

Figure 1. HIV transmission and mortality by CD4 count. (A) HIV transmission rate per 100 person-years (PYs) by CD4 count for the infected partner in discordant couples enrolled in a randomized controlled trial of acyclovir [11]. (B) Mortality rate by CD4 category in ART-naïve HIV-positive individuals enrolled in research cohorts in West Africa [132]. In both panels, the width of the bars represents the proportion of ART-naïve HIV-positive 15- to 64-year-olds by CD4 count in a nationally representative household survey in Kenya [15].
doi:10.1371/journal.pmed.1001258.g001

Another testing-related problem is that within-patient variability in CD4 cell count can be very high [19], such that the CD4 count from a single test could be an unreliable indicator of transmission risk and clinical need [20,21]. Moreover, HIV-infected individuals who are feeling healthy may decline the option to initiate treatment [22], a challenge likely to be exacerbated under earlier treatment eligibility at high CD4 counts. Earlier access to ART would, on the other hand, also reduce the number of patients needing pre-ART care, the phase at which retention is the poorest, according to a systematic review of retention in HIV care in sub-Saharan Africa [17].

Steadily increasing the CD4 threshold for treatment eligibility as further resources become available may be viewed as the most equitable and acceptable strategy for allocating additional treatment, considering that treatment eligibility has long been based on a CD4 criterion, but while resources for treatment continue to be constrained, expanding treatment access beyond current clinical guidelines based on an increasing CD4 criterion is unlikely to be the most efficient route to maximising the epidemiological or clinical benefit of ART.

Viral Load

Untreated asymptomatic HIV infection is characterised by the viral load fluctuating around a steady level, called the set-point viral load (SPVL) [23]. Individuals vary considerably in SPVL; values are approximately log-normally distributed with standard deviation 0.75 $\log_{10}$ units, such that the 95% range spans a 1,000-fold variation in SPVL [24]. SPVL has proven one of the more robust predictors of infectiousness [11,25–28]. In a recent study amongst serodiscordant couples [28], the transmission rate in couples with index individuals with viral load in the range 100,000 to 1,000,000 copies/ml of blood was estimated to be 5.6 per 100 person-years at risk (95% confidence interval: 4.0 to 7.6), while the transmission rate for index individuals with viral load in the range 100 to 1,000 copies/ml was estimated at 0.8 per 100 person-years (0.4 to 1.5). Thus, a 100-fold difference in viral load translates to a 7-fold variation in infectiousness, although this relationship is highly nonlinear [11,25–28]. As it becomes easier to measure viral loads in the field, with point-of-care tests in development (e.g., [29]), it becomes reasonable to ask whether prioritising further ART expansion for individuals with high viral load would be an effective, efficient, and affordable strategy.

While individuals with high SPVL are more likely to effectively transmit the virus, they also tend to progress from asymptomatic infection to disease more quickly than those with low SPVL [30]. To estimate how much individuals with differing SPVL contribute to the epidemic, their transmission potential can be calculated as the product of their biological infectiousness and duration of

infection [24] (Figure 2). Compared to individuals with intermediate SPVL, individuals with very high SPVL may contribute less to the epidemic, because they progress to advanced disease and death very quickly and thus have fewer opportunities to infect others.

Consequently, prioritised ART expansion for individuals with very high viral loads may not provide greater long-term prevention benefits than expanded access for a comparably large random fraction of the untreated population. The principal frailty in this conclusion comes from multiplying infectiousness and duration of asymptomatic infection from different studies. However, this conclusion is robust to parametric assumptions, to assumptions about the sexual network, and to including heightened infectiousness in early- and late-stage untreated infection [24,31].

While the epidemiological benefit of expanded ART access for individuals with very high SPVL may be limited, targeting these individuals for rapid ART initiation may offer substantial clinical benefits. ART prioritisation for people with very high SPVL could be expensive to implement, as viral load screening and follow-up would require substantial resources. How this form of prioritisation would affect the number of patients eligible for treatment is not clear: a recent analysis of HIV-1 RNA viral load data from two general population cohorts in Botswana suggested that 24%–28% and 14%–18% of HIV-infected, treatment-naïve individuals ($n = 1,286$) had viral load levels greater than 50,000 and 100,000 copies/ml, respectively [32], but it is unclear how many of these individuals were not eligible under current CD4-based ART initiation guidelines.

Pregnant Women

Existing guidelines for the prevention of mother-to-child transmission (PMTCT) recommend that pregnant women with CD4 counts higher than 350 cells/μl take an antiretroviral drug course from the 14th week of pregnancy until one week after delivery (Option A) or until one week after breastfeeding has finished (Option B) [33]. A new option "B+" has been proposed, in which pregnant women would be eligible to immediately initiate lifelong ART regardless of HIV disease stage, TB disease status, or CD4 count [34]. The cost and epidemiological impact of expanding ART to all pregnant women will vary between settings with different patterns of fertility, sexual behaviours, and existing ART programmes.

The potential HIV prevention impact of option B+, beyond PMTCT, would be low if many infected pregnant woman are in stable relationships with partners who are already infected. Data from Demographic and Health Surveys in Lesotho, Malawi, and Kenya indicate that more than half of married, cohabiting partners of HIV-infected pregnant women are HIV-infected (83% [10/12] in Lesotho, 54% [20/37] in Malawi, and 50% [6/12] in Kenya) [35–37]. However, for serodiscordant couples, the female-to-male transmission rate may be more than twice as high during pregnancy as during non-pregnant periods [38]. Whether ART initiation during pregnancy would effectively override this risk elevation is questionable, given the lag time of up to five months between ART initiation and viral load suppression [39].

The number of additional people who would be on treatment with this prioritisation strategy depends on several factors. The crude birth rate (and hence the incidence of pregnancy) varies greatly across sub-Saharan Africa, even within subregions: from as high as 46.5 childbirths per 1,000 people per year in Zambia to less than half this rate (22.9/1,000 individuals/year) in the neighbouring country of Botswana [40]. However, overall, the difference between the number of HIV-infected pregnant women that would be ART-eligible under PMTCT option B+ and the number eligible under current ART initiation guidelines may be small because of the effect of haemodilution on CD4 cell count. Haemodilution, a normal physiological phenomenon during pregnancy, temporarily reduces CD4 cell count, meaning that many pregnant HIV-infected women become eligible for treatment on the basis of CD4 count during pregnancy. In a prospective cohort study of 324 HIV-infected pregnant women from Abidjan, Cote d'Ivoire, 48.3% (157/325) had CD4 counts less than 350 CD4 cells/μl at 32 weeks of gestation, yet this fraction decreased to 28.9% (94/325) one month postpartum [41]. The implication for Cote d'Ivoire is that only just over 10,000 additional women would initiate ART if immediate treatment was expanded to all pregnant women regardless of CD4 count, and ART coverage in women, in the first year of the intervention, would increase only from 39% to 42% [42].

HIV-positive pregnant women are a priority group that is relatively easy to identify, because of the high uptake of antenatal care (ANC) in most populations, with associated HIV counselling and testing, even in resource-limited settings. Several studies reported very high acceptance of provider-initiated HIV counselling and testing in ANC at several sites across Africa during the past few years: 99.5% acceptance of testing in Nigeria, 91% in South Africa, 97% in Ghana, and 99% in Zambia [43–46]. The acceptance of HIV testing at first ANC visit is still as low as 69.1% in rural areas of Swaziland and South Africa [47], but provider-initiated testing and counselling in ANC may be able to raise the testing uptake by 9.9%–65.6% [48].

Obstacles remain in the linkage between diagnosis in ANC and long-term ART treatment because of ART refusal and poor retention. In a recent review, Ferguson et al. found that 38%–88% of known ART-eligible women in sub-Saharan countries fail to initiate treatment [49]. Once in treatment, retention among pregnant women has been found to be no worse than in other population groups in seven resource-limited countries in sub-Saharan Africa and Thailand [50]. However, Boyles et al. found that initiating ART while pregnant is associated with a higher lost-to-follow-up risk compared with the general population in rural South Africa [51]. Retention challenges faced when expanding ART to pregnant women regardless of CD4 count are likely to be similar to those currently faced in traditional PMTCT programmes: (1) patients' not being prepared for HIV testing and its implications before the ANC visit; (2) fear of stigma, discrimination, household conflict, or divorce on disclosure of HIV status; (3) long waiting times at the ANC facilities; and (4) inability to afford the transport to these facilities [52].

Because expanding access to ART for pregnant women utilises existing ANC and PMTCT infrastructure for diagnosis and HIV counselling, the only additional costs associated with this strategy are additional drug costs for the period between the end of pregnancy and ART eligibility under other criteria, suggesting favourable affordability of this ART expansion strategy. Cost-effectiveness studies of ART in pregnant women have thus far focused on benefits in terms of PMTCT, and have found that it is cost-effective as measured against accepted international benchmarks in a variety of low- and middle-income countries [53,54]. Cost-effectiveness studies of PMTCT option B+ for adult HIV transmission prevention are still to be conducted. Expanding ART to all HIV-positive pregnant women may provide additional maternal health benefits and contribute to the Millennium Development Goals if ART, PMTCT, and reproductive health care services are integrated [55,56].

Active Tuberculosis Disease

The provision of ART to all HIV-infected people with active TB disease, irrespective of CD4 cell count, has been recom-

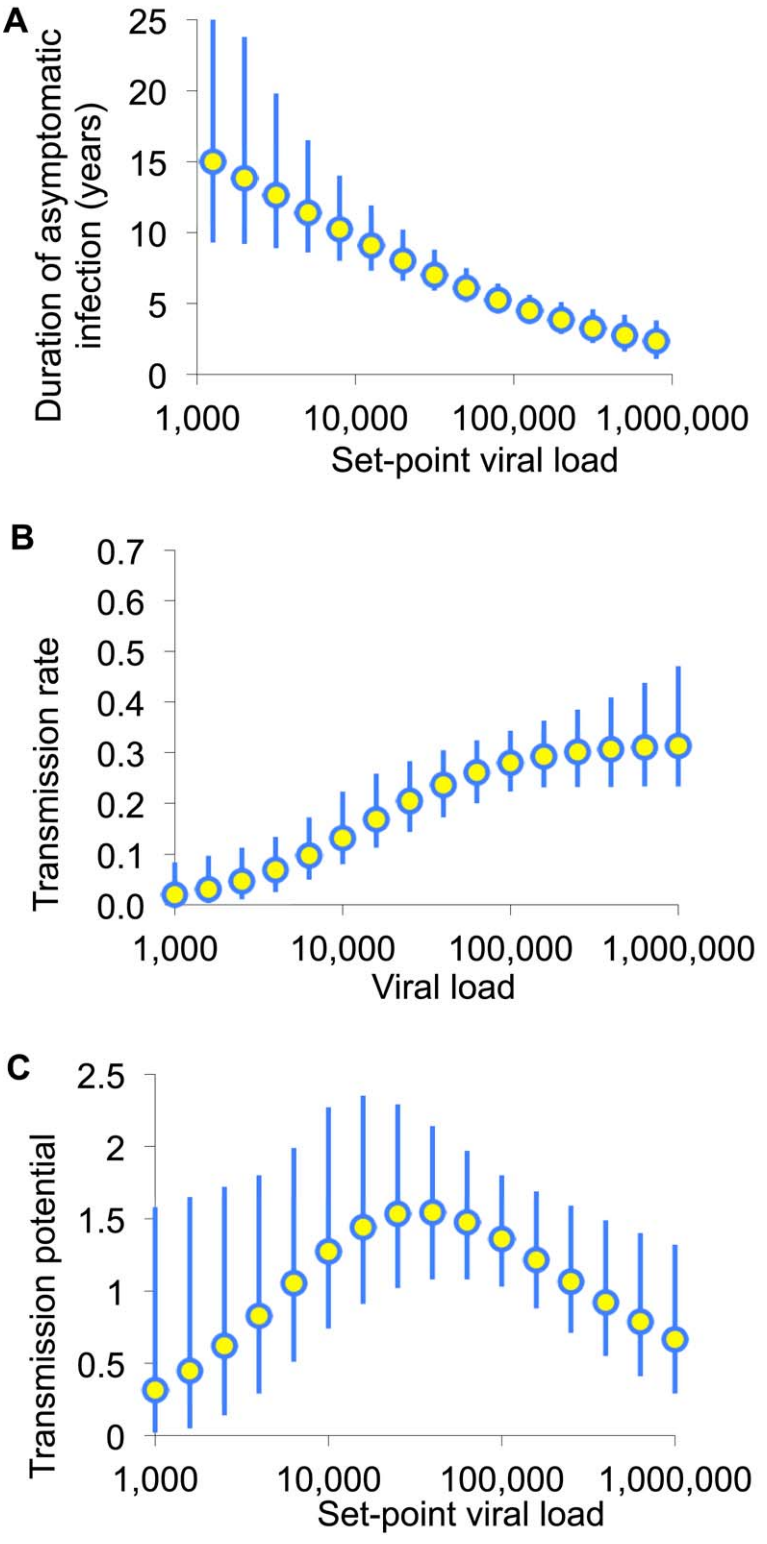

Figure 2. The transmission potential of individuals as a function of set-point viral load. (A) Infectiousness (per unit calendar time) and (B) duration of asymptomatic infection are estimated by fitting to various sources of data as described in [24]. (C) The product of these is the transmission potential, the average number of people an infected individual is expected to infect over the whole of asymptomatic infection. The transmission potential measures the relative prevention effect of treatment as prevention targeted to an individual with a given SPVL. Adapted from Fraser et al. [24].
doi:10.1371/journal.pmed.1001258.g002

mended by the World Health Organization since 2010 [3], based on its clinical benefits. In the SAPiT trial, the mortality rate in 429 patients with CD4 cell counts up to 500 cells/μl who initiated ART during TB treatment was 56% lower (95% confidence interval: 21%–75%, $p = 0.003$) than in patients who initiated ART after completion of TB treatment [57]. However, the coverage of ART for all HIV-infected people with active TB disease remains low. In the WHO African Region in 2011, only 59% of TB patients were tested for HIV, and of those identified to be HIV-infected, only 42% were on, or started on, ART [58].

The epidemiological benefits of expanding ART to all patients with active TB disease are unclear. Although there were some early indications that those with TB disease are more infectious [59–61], the largest study conducted among HIV-positive people with incident TB disease indicated that viral load increases by only a small amount following a TB episode [62], and a more recent study showed that treating active TB disease in individuals with CD4 counts greater than 350 cells/μl reduced markers of immune activation but had no effect on HIV viral load or CD4 count [63]. Therefore, providing ART to TB patients with CD4 counts above 350 cells/μl is likely to have a similar prevention effect on HIV transmission as treating a random subset of HIV-infected individuals with a CD4 count above 350 cells/μl.

In high HIV prevalence settings, the proportion of HIV-infected people with active TB who have CD4 cell counts greater than 350 cells/μl has been estimated to range from 11% to 30% [2]. Assuming a 1% incidence of active TB disease and 50% HIV prevalence in individuals with incident active TB, this would mean that for South Africa, around 27,500 to 75,000 extra individuals would be eligible for ART in the first year of this form of prioritisation. Given the suppressive effect of ART on TB disease incidence [64,65], a decreasing number of active TB patients in need of ART would be expected in the following years. A modelling study that estimated the impact of the roll-out of annual HIV testing and immediate ART on TB disease incidence in nine African countries reported a 21% (range: 10%–31%) reduction in the cumulative AIDS-related TB disease incidence over the first five years, and a 48% (range: 37%–55%) reduction in the incidence of TB disease at five years [64].

Integration of ART provision for all HIV patients, regardless of TB coinfection status, with TB services may offer a relatively feasible way to implement an expansion of ART to individuals with active TB disease. Data from eight countries with a high burden of HIV-infection-associated TB showed that there were up to five TB treatment facilities for each ART facility in 2007 [66], and a study in Tugela Ferry, South Africa, showed that integration of TB and HIV services was associated with high ART adherence [67]. However, with this approach it would be critical to implement adequate infection control to minimise nosocomial TB infection, and obtaining high TB treatment coverage is challenged by the difficulty of diagnosing active TB in HIV-infected patients [68].

Given the clear clinical benefit of ART in TB patients, this option of ART expansion is likely to be highly acceptable by both the target group and the general population. For TB patients, current illness and the prospect of a reduced risk of TB recurrence are incentives for ART initiation, high adherence, and retention in care. If implemented successfully, ART expansion to all TB patients should lead to large gains in healthy person-years of life and long-term cost savings due to decreased recurrent TB.

Serodiscordant Long-Term Relationships

Stable serodiscordant relationships, in which one partner is HIV-infected and the other is not, are an identifiable prevention opportunity, and the continued transmission in such couples during carefully monitored clinical trials with intensive counselling demonstrates the need for additional prevention options for this population [5,11]. Trial and observational data have demonstrated the efficacy of ART in preventing HIV transmission in stable serodiscordant heterosexual partnerships [5,11], and recent WHO guidelines for stable serodiscordant couples already include offering ART to the HIV-infected partner irrespective of CD4 cell count, in addition to behaviour change counselling [69]. While the biological efficacy of the effect of ART on transmission risk should generalise to non-stable heterosexual partnerships as well, it has been hypothesized that couples in stable partnerships will be most able to adhere to daily dosing regimens and therefore achieve the maximum individual-level benefit [70,71]. Further, it is known that couples in stable partnerships in which the HIV-infected individual has a high CD4 cell count are likely to conceive (16% per year among discordant couples [38]); therefore this strategy would incur many of the maternal and PMTCT-related health benefits described above [33].

The relative epidemiological impact of prioritising early treatment to HIV-infected individuals who have an uninfected long-term partner will depend primarily on the risk of within-couple transmission without treatment, and secondarily, on the risk of onward transmission from the partner to someone else. The risk of transmission without ART in couples could be relatively low: 1.7 per 100 person-years at risk ([5]; among those with CD4 counts of 350–500 cells/μl: [11]). One model suggests that providing ART to serodiscordant couples might be expected to avert 21 infections per 1,000 person-years of ART [72]. If the risk of transmission in couples is actually higher (as has been observed in couples that did not necessarily know that they were in a discordant partnership [73], and as assumed by El-Sadr et al. [74]), and if it is assumed that the infected partner forms many additional partnerships with other individuals, then it has been estimated that the number of infections averted per 1,000 person-years of ART could be as high as 77. This can be compared to 53–159 for providing ART to all individuals with CD4 cell counts below 350 cells/μl irrespective of partnership status, and 65–152 if ART is provided irrespective of CD4 cell count [75]. Thus, prioritising those in stable partnerships for treatment may not be a more efficient form of prevention than providing the treatment to the general population without prioritisation.

It is unclear how feasible it would be to preferentially provide access to ART to those in serodiscordant couples. Only ~8%–31% of couples were found to be discordant in recruitment to a clinical trial [76], and other data suggest that the countries with the highest levels of HIV prevalence tend to have the smallest numbers of stable serodiscordant couples [77]: only a small minority (<15%) of infected individuals report being in a stable partnership with someone known to be uninfected [77]. There are few opportunities to identify serodiscordant couples in current health care systems in most settings in Africa, though household testing interventions may increase opportunities to reach couples [78]. However, many would question the general acceptability of an intervention that favours those in stable discordant partnerships over those in concordant partnerships. Operationally, defining a consistent criterion for a discordant couple is challenging. In Kenya, for example, there may be 150,000 individuals that would be newly eligible to start treatment today under this policy [77], but many more might claim to be in stable discordant relationships, or limitations in disclosure in couples could mean that many fewer would actually start treatment earlier.

Female Sex Workers

Almost one-fifth of the HIV epidemics in sub-Saharan Africa are classified as concentrated (defined as HIV epidemics with HIV

prevalence < 1% in the adult population), and many more are not highly disseminated (43% of epidemics in this region have an adult HIV prevalence below 3%) [79]. In these settings, FSWs and their clients are key populations for the transmission of HIV [80–86]. Previous modelling [81,82,87,88] and epidemiological analyses [89,90] suggest that prioritising interventions for FSWs and their clients in these settings can substantially reduce HIV transmission amongst FSWs, and amongst the population as a whole. It therefore seems natural to consider whether ART eligibility irrespective of CD4 cell count should be prioritised to FSWs.

A literature review was conducted in PubMed with the search terms "("sex workers" or FSW or FSWs or CSW [commercial sex worker] or CSWs or sexwork*) and HIV and (antiretroviral or "anti-retroviral" or ART or HAART)". This produced 67 papers, of which nine considered ART treatment amongst FSWs, barriers to accessing care, and risk behaviours following ART initiation [91–97]. The majority were from sub-Saharan Africa (six of nine), with three others from Chennai, India, and Vancouver, Canada. These papers informed the following discussion of ART prioritisation for FSWs.

In settings with existing and effective non-ART interventions to prevent HIV transmission amongst FSWs and their clients, the main questions are to evaluate the potential added prevention benefits of prioritising ART to FSWs, and the likelihood of risk substitution (i.e., potential increases in risk behaviours following ART expansion) [94,98,99]. In settings where behaviour-targeted interventions have not been fully implemented, the question is whether these should be scaled up before scaling up ART for prevention. If expanded access to ART is scaled up for FSWs following behaviour change interventions, this may increase the relative impact of ART treatment. For instance, because increases in condom use could reduce the incidence of new acute HIV infections, it is possible that such an intervention could temporarily lead to a smaller proportion of incident infections being due to early acute infection, and therefore a relatively greater prevention benefit of ART when provided to those with chronic infection.

Achieving a high preventive benefit from expanding ART to FSWs depends on initiating and retaining individuals in programmes. FSWs have generally received lower coverage of ART, because of factors such as reduced health-seeking behaviour and the stigmatised nature of sex work [100]. However, numerous targeted HIV prevention interventions worldwide show that FSWs can be engaged and recruited into intensive interventions with high coverage [101–103] and at reasonable cost [104–106]. Emerging data on ART provision amongst FSWs [92,93,95,107] suggest that FSWs can be successfully initiated on ART in resource-poor settings; existing interventions could act as an easy and affordable entry point for increasing ART coverage among FSWs [93]. However, maintaining high ART adherence among FSWs remains challenging, leading to poorer outcomes with respect to CD4 count and suppression of viral load compared to non-FSWs [93,107]. This is likely to translate into smaller reductions in infectivity, and greater morbidity or mortality [93,107,108], and indicates that there would be a particular need for retention efforts and adherence counselling for this prioritisation group [96], which could increase the costs of FSW-targeted ART programmes.

In addition to clinical and behavioural issues, the transient nature of sex work could affect the potential impact of ART on transmission, and the subsequent costs. As most FSWs sell sex for only a few years [109,110], the early phase of acute HIV infection with high HIV viraemia may make a disproportionately large contribution to sex-work-driven HIV transmission [111]. Even with frequent retesting and immediate linkage to care, ART is likely to be initiated only after this phase, which would reduce the impact of ART on transmission, and highlights the continued importance of condom interventions. In addition, the cumulative costs of immediate ART eligibility for FSWs might grow sharply, as new FSWs become infected and eligible for ART, and HIV-infected former FSWs remain on ART after ceasing sex work. This could result in considerable investment and a suboptimal allocation of ART in some settings with rapid turnover of sex worker populations. However, it is possible that ex-FSWs may still be at greater risk of transmitting HIV than the general population and so could remain a good target population for ART as prevention.

Lastly, as has been shown for other HIV prevention interventions, the expected preventive benefit of targeting FSWs with ART is smaller in generalised epidemics than in concentrated epidemics, and is reduced in the late phase of an HIV epidemic compared to the early phase, especially in the short term. However, even in generalised HIV epidemics, modelling analyses suggest it would be cost-effective to target FSWs because of their disproportionate contribution to HIV transmission, although it may not be sufficient for achieving large and rapid reductions in HIV transmission in the general population. Conversely, not reaching high-risk groups such as FSWs may seriously attenuate the impact of any ART intervention [71,112].

In summary, the decision to target FSWs with ART provision has to balance the likely heightened costs associated with increased adherence counselling and monitoring, and outreach to ensure retention, with benefits of decreased transmission that may be short-lived in contexts where sex work is transient. However, in settings where sex work is longer term the impact could be much greater. In addition, the ethical and social acceptability of giving prioritised ART access to FSWs needs to be carefully considered before any FSW-targeted programme is initiated—the benefits to the population as a whole would need to be clearly determined and communicated, and proactive monitoring of all ART provision channels would need to be in place to ensure that the care of other HIV-infected individuals is not compromised. Drug resistance should also be monitored, as FSWs on ART may facilitate the spread of resistance.

Men Who Have Sex with Men and People Who Inject Drugs

The arguments for expanded access to other key populations, including MSM and PWID, are similar to those for expanded access to FSWs. If there is a population that contributes disproportionately to the number of infections in a population, and they can be identified and enrolled and retained in care, then it could be efficient to prioritise ART access to that group. However, the evidence from the HPTN 052 study that ART reduces infectiousness was specifically for heterosexual transmission: the extent to which ART decreases transmission occurring through homosexual sex or intravenous injection is not known [113].

Also, for the epidemics in Africa, there is little information about the population sizes of MSM and PWID, and their behaviours and contribution to the epidemic, which makes it hard to formulate firm recommendations about the benefits of prioritising access to these groups. Several studies in Africa have been able to recruit MSM [114,115], and it has been estimated that, in total, transmission among MSM could account for 6% of new infections in Kenya and up to 21% in some concentrated epidemics [115], a range that is broadly supported by the Joint United Nations Programme on HIV/AIDS review of modes of transmission ([116]; K. K. Case, P. D. Ghys, E. Gouws, J. W.

Eaton, P. Cuchi, et al., unpublished data]. Meanwhile, in a global review of injecting behaviour, there were no data (or estimated prevalence levels) for most African countries [117], and consequently the estimated number of HIV-infected PWID was very uncertain, ranging from 26,000 to 572,000 for sub-Saharan Africa. However, it has been hypothesized that in particular areas, such as Mombasa and Nairobi in Kenya, a high frequency of injecting among a growing population of PWID, coupled with overlapping sexual risk behaviours, has resulted in a substantial proportion of overall transmission possibly resulting from injection [118,119]. There are also indications that PWID are less likely to access care and treatment services than others [120], and they have lower adherence [121] and retention to therapy [122,123], so any ART programme prioritising ART to this group would presumably have to contend with these issues.

Discussion

If it could be afforded, all HIV-infected individuals who wanted to initiate ART should be able to do so. However, resource constraints, at least in the short and medium term, necessitate some form of prioritising of HIV treatment through health policies. These policies should maximise epidemiological and clinical benefit while still being feasible, affordable, acceptable, and equitable [124]. To date, this prioritisation has been based principally on the CD4 cell count of HIV-infected individuals, as a marker of their immediate clinical need, but with the finding that

Table 1. Likely profile of prevention and clinical impact, affordability, feasibility, and acceptability of alternative options for ART expansion beyond current guidelines.

Prioritisation Group	Impact on New HIV Infections	Impact on HIV-Related Morbidity and Mortality	Feasibility	Affordability	Acceptability
CD4 cell count (350–500 cells/μl)	− (Unlikely to be highly transmissible, relative to those at lower CD4 cell counts or other prioritisation groups)	? (Clinical trial evidence expected from START trial, reporting in 2015; unlikely to be as efficient as strategies targeting clinical need, e.g., high SPVL, TB coinfection)	+ (Screening utilises already standard CD4 screening; reaching high coverage would likely require efforts to improve routine HIV testing at the population level)	− (Would likely expand access to treatment to an additional 20% of the HIV-infected populations)	+ (May be perceived as the most equitable option for expanding access to ART, because of the history of determining treatment need and access based on CD4 cell count)
Viral load (SPVL≥50,000 copies/ml)	• (Strong evidence from many discordant studies that infectiousness increases with SPVL, but not dramatically)	+ (Strong evidence from many seroconverter cohorts that individuals with high SPVL progress rapidly to AIDS, and so may enhance linkage to care in rapid progressors)	? (Requires development of point-of-care viral load testing; many prototypes, but none validated yet)	? (Cost of point-of-care viral load testing is unknown)	? (May prove controversial if not backed by evidence for direct clinical benefit)
Active TB disease	− (Likely to have the same impact on HIV transmission as reaching a subset of HIV-infected individuals with CD4 cell counts between 350 and 500 cells/μl)	+ (Much greater impact on morbidity and mortality than treating many other groups)	+ (Can be integrated with existing TB services; adherence/retention to ART may be higher because of current illness and the prospect of a reduced risk of TB recurrence)	+ (Relatively small group, compared with individuals with CD4 350–500 cells/μl; large reduction in mortality suggests targeting TB patients may be more cost-effective than other groups)	+ (Given the clear clinical need, likely to be highly acceptable to both the target group and the general population)
Pregnant women	? (Potential reductions in maternal orphanhood and potential to prevent paediatric HIV transmission; estimates of the impact on heterosexual HIV transmission are yet to be produced)	+ (Impact mainly on morbidity and mortality of newborns with HIV-positive mothers)	+ (Targets are easy to identify via existing ANC; testing uptake is high in some areas and can be increased by provider-initiated service; contrary results are found on retention)	+ (Increment of newly identified target patients is not big; infrastructures and staff that already exist favours the affordability)	+ (May be better accepted by patients if initiated by ANC provider)
Serodiscordant couples	− (Likely fewer infections averted per person-year of ART than allocation to those with multiple partners)	? (Marginal therapeutic benefit of ART initiation at CD4 >350 cells/μl not certain)	+ (In some settings couples hard to find; trial data indicate discordant couples are a highly motivated population with good adherence to pill-taking regimes and retention in care)	? (Minority of infected individuals in stable discordant couples, but uptake unknown)	? (Unclear if it is socially acceptable for those with stable partners to receive preferential access)
Sex workers	+ (Elevated HIV transmission risk in many settings likely to result in large number of HIV infections averted per year on ART)	? (May be more modest than other groups because limited data suggest that they have lower adherence and worse outcomes in terms of morbidity and mortality)	+ (Previous FSW-targeted interventions have demonstrated feasibility; limited studies suggest FSWs are willing to initiate ART; however, likely to have worse adherence and retention)	+ (FSWs make up a small proportion of the female population, but if sex work is of short duration, then there may be a much larger group of ex-FSWs that will continue on ART)	? (May not be acceptable to the wider community; programmes would need to show and emphasise population benefit)

?, insufficient evidence to warrant definitive decision; +, available evidence suggests this is a beneficial option, compared to the other expansion options; −, available evidence suggests this is an unfavourable option, compared to the other expansion options; •, available evidence supports neither that this is a beneficial option nor that it is an unfavourable option.

doi:10.1371/journal.pmed.1001258.t001

ART reduces transmission risk, it is important to re-evaluate other ways in which ART could be allocated. In this article, we have examined several of the main options for prioritising ART access and have highlighted the key epidemiological and policy considerations that should guide decision-making and future research (summarised in Table 1).

There are some forms of prioritisation that are already supported by existing guidelines or programmes. In particular, ART for all individuals with active TB disease has substantial epidemiological and clinical benefits, and already forms part of WHO international guidelines. Treatment for pregnant women irrespective of CD4 cell count, for which the epidemiological impact is not yet clear, could have advantages in terms of simplified care for pregnant women and benefits for their children and partners, and is being implemented in some settings.

Important questions remain regarding all of the options, and there is a clear need for further data collection. Some knowledge gaps could be filled shortly, as results are reported from at least 50 projects planned or ongoing to evaluate the impact of ART and other interventions on HIV- and/or TB-related morbidity and mortality, HIV incidence and transmission, and risk behaviour [125]. Almost half of these projects are in sub-Saharan Africa, and include studies that will test the individual- and community-level preventive effect of ART provided to patients with CD4 cell counts between 350 and 500 cells/μl, those with the highest viral loads, HIV-infected pregnant women, and HIV-infected partners in serodiscordant couples [125,126]. In addition, the secondary objectives of many of these projects are the evaluation of the feasibility, cost, health care impact, treatment adherence and retention, and social acceptability of the interventions. With a large variation in geographical areas, target populations, and outcome variables, the combined body of evidence generated by these studies may begin to address the question of whether and how different sociodemographic, economic, and epidemiological contexts influence the impact of ART interventions.

In the short term, the costs of expanding access to ART are likely to be driven by the size of the groups to whom access is extended and the costs associated with identifying and reaching members of these groups. Long-term affordability is likely to depend on the size of the group as well as reductions in incidence resulting from the expanded ART programme, the success of other HIV prevention interventions, and economic nonlinearities such as economies of scale. Although relative group sizes will vary from setting to setting, ART expansion to HIV-infected people with CD4 counts of 350–500 cells/μl or above 500 cells/μl would likely require the largest programme increase. In contrast, initial increases required to prioritise ART in FSWs and pregnant women would most likely be much smaller than for the other prioritisation options, though cumulative costs would grow as women started ART during pregnancy or sex work, then continued on lifelong ART.

Evidence from ART programmes in southern Africa indicates that high retention in care becomes increasingly challenging as treatment programmes expand [127]. Affordability and feasibility are negatively affected not only by larger group size, but also by the more intensive efforts required to identify eligible people and maintain high adherence and retention in care. Globally, patients' health literacy regarding ART adherence remains an important challenge [128]. On the other hand, there is a rapidly growing body of strategies and tools to improve retention in care and ART adherence, including interventions to improve the mental health (especially treatment of depression) of HIV-positive individuals, and reminder devices and interactive communication technologies [129].

This review has aimed to highlight some of the key issues and identify the needs for future studies, and has not provided a direct quantitative comparison of the impact of alternative prioritisation strategies in specific settings, which will be a critically important body of modelling work in the future. To further facilitate a constructive debate that is meaningful to national decision-makers and donor organisations, context-specific mathematical models should be developed that enable head-to-head comparison of multiple ART expansion options in an internally consistent manner, that is, with all simulations based on the same data and assumptions. However, the considerations raised here already indicate that the impact and feasibility of these alternative forms of ART allocation are expected to vary substantially between settings, and there is no single formulation that will be optimal in all settings. Furthermore, the best strategy will depend on the relative values assigned to therapeutic benefits, preventative benefits, and wider societal benefits, such as reducing the number of orphans and increasing labour force availability. Combined metrics of impact such as quality-adjusted life years saved or disability-adjusted life years averted [130] can be used to understand how preventative and therapeutic benefits are related.

The overall effectiveness of treatment in reducing infectiousness, as well as the risk of drug resistance [130], is expected to be crucially dependent on the viral suppression achieved, which is in turn affected by patterns of adherence. Throughout this analysis, we have assumed ART has a suppressive effect on HIV transmission for all patients receiving the treatment. Whilst this is biologically plausible, we recognise that it is possible that different groups could behave differently from the HIV-infected individuals in the HPTN 052 trial and therefore achieve lower levels of viral suppression and a smaller reduction in infectiousness. However, there is little information available on levels of viral suppression for ART, nor on adherence to the treatment regimen.

Finally, we recognise that the issues of expanding access to ART do not exist in a vacuum. Decisions concerning the implementation of ART should be scaled—scaling up will have to take place in the context of the entire portfolio of the HIV response programme in a particular country, which will include multiple forms of prevention intervention. Indeed, WHO guidance on the use of antiretrovirals for prevention is expected to include both pre-exposure prophylaxis and ART, and we would anticipate that further strategic advice from normative agencies will increasingly

Key Points

- Discussions about expanded access to ART for HIV prevention have been focused on one particular strategy—providing ART to all HIV-infected individuals. Here we aim to broaden the discussion by considering the implications of prioritising access to ART according to clinical and behavioural factors.
- Any recommendation to prioritise particular groups should consider not only the impact of ART in that group, including its therapeutic and prevention effects, but also its feasibility, affordability, and acceptability.
- Some forms of prioritisation—ART for individuals with active TB and for pregnant women irrespective of CD4 cell count—are already promoted by existing guidelines or programmes.
- For other prioritisation options, there are currently insufficient data to make first recommendations, although findings of future studies and further modelling analyses should contribute to forming policy.

embrace the full range of possibilities for maximising the health impact of ART in combination with other interventions [131].

Author Contributions

Wrote the first draft of the manuscript: WD JWE FM CF RGW PV M-CB TBH. Contributed to the writing of the manuscript: WD JWE FM CF RGW PV M-CB TBH. ICMJE criteria for authorship read and met: WD JWE FM CF RGW PV M-CB TBH. Agree with manuscript results and conclusions: WD JWE FM CF RGW PV M-CB TBH.

References

1. World Health Organization, Joint United Nations Programme on HIV/AIDS, United Nations Children's Fund (2011) Global HIV/AIDS response: epidemic update and health sector progress towards universal access: progress report 2011. Geneva: World Health Organization. Available: http://whqlibdoc.who.int/publications/2011/9789241502986_eng.pdf. Accessed 13 December 2011.
2. Sadr WE, Cohen M, DeCock K, Bekker LG, Karim S, et al. (2011) PEPFAR Scientific Advisory Board recommendations for the Office of the US Global AIDS Coordinator: implications of HPTN 052 for PEPFAR's treatment programs. Washington (District of Columbia): Office of the US Global AIDS Coordinator.
3. World Health Organization (2010) Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach. 2010 revision. Geneva: World Health Organization.
4. Cohen J (2011) Breakthrough of the year. HIV treatment as prevention. Science 334: 1628.
5. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
6. Hargrove JW, Humphrey JH (2010) Mortality among HIV-positive postpartum women with high CD4 cell counts in Zimbabwe. AIDS 24: F11–F14.
7. (2010) Funding, other trends imperil global response to HIV/AIDS. AIDS Policy Law 25: 4.
8. Geng EH, Bwana MB, Kabakyenga J, Muyindike W, Emenyonu NI, et al. (2010) Diminishing availability of publicly funded slots for antiretroviral initiation among HIV-infected ART-eligible patients in Uganda. PLoS ONE 5: e14098. doi:10.1371/journal.pone.0014098
9. Panel on Antiretroviral Guidelines for Adults and Adolescents (2011) Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Washington (District of Columbia): US Department of Health and Human Services. Available: http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. Accessed 4 January 2012.
10. European AIDS Clinical Society (2011) Guidelines: version 6—October 2011. Available: http://www.europeanaidsclinicalsociety.org/images/stories/EACS-Pdf/EACSGuidelines-v6.0-English.pdf. Accessed 12 June 2012.
11. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, et al. (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet 375: 2092–2098.
12. Sterne JA, May M, Costagliola D, de Wolf F, Phillips AN, et al. (2009) Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. Lancet 373: 1352–1363.
13. Van Cutsem G, Ford N, Hildebrand K, Goemaere E, Mathee S, et al. (2011) Correcting for mortality among patients lost to follow up on antiretroviral therapy in South Africa: a cohort analysis. PLoS ONE 6: e14684. doi:10.1371/journal.pone.0014684
14. ClinicalTrials.gov (2012) Strategic Timing of Antiretroviral Treatment (START). Available: http://clinicaltrials.gov/ct2/show/NCT00867048. Accessed 4 January 2012.
15. Stover J, Bollinger L, Avila C (2011) Estimating the impact and cost of the WHO 2010 recommendations for antiretroviral therapy. AIDS Res Treat 2011: 738271.
16. Joint United Nations Programme on HIV/AIDS (2011) UNAIDS world AIDS day report 2011: how to get to zero. faster. smarter. better. Geneva: Joint United Nations Programme on HIV/AIDS.
17. Rosen S, Fox MP (2011) Retention in HIV care between testing and treatment in sub-Saharan Africa: a systematic review. PLoS Med 8: e1001056. doi:10.1371/journal.pmed.1001056
18. Kranzer K, Zeinecker J, Ginsberg P, Orrell C, Kalawe NN, et al. (2010) Linkage to HIV care and antiretroviral therapy in Cape Town, South Africa. PLoS ONE 5: e13801. doi:10.1371/journal.pone.0013801
19. Novitsky V, Ndung'u T, Wang R, Bussmann H, Chonco F, et al. (2011) Extended high viremics: a substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. AIDS 25: 1515–1522.
20. Malone JL, Simms TE, Gray GC, Wagner KF, Burge JR, et al. (1990) Sources of variability in repeated T-helper lymphocyte counts from human immunodeficiency virus type 1-infected patients: total lymphocyte count fluctuations and diurnal cycle are important. J Acquir Immune Defic Syndr 3: 144–151.
21. Raboud JM, Haley L, Montaner JS, Murphy C, Januszewska M, et al. (1995) Quantification of the variation due to laboratory and physiologic sources in CD4 lymphocyte counts of clinically stable HIV-infected individuals. J Acquir Immune Defic Synd 10 (Suppl 2): S67–S73.
22. Katz IT, Essien T, Marinda ET, Gray GE, Bangsberg DR, et al. (2011) Antiretroviral therapy refusal among newly diagnosed HIV-infected adults. AIDS 25: 2177–2181.
23. Geskus RB, Prins M, Hubert JB, Miedema F, Berkhout B, et al. (2007) The HIV RNA setpoint theory revisited. Retrovirology 4: 65.
24. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP (2007) Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. Proc Natl Acad Sci U S A 104: 17441–17446.
25. Baeten JM, Kahle E, Lingappa JR, Coombs RW, Delany-Moretlwe S, et al. (2011) Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. Sci Transl Med 3: 77ra29.
26. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med 342: 921–929.
27. Fideli US, Allen SA, Musonda R, Trask S, Hahn BH, et al. (2001) Virologic and immunologic determinants of heterosexual transmission of human immunodeficiency virus type 1 in Africa. AIDS Res Hum Retroviruses 17: 901–910.
28. Lingappa JR, Hughes JP, Wang RS, Baeten JM, Celum C, et al. (2010) Estimating the impact of plasma HIV-1 RNA reductions on heterosexual HIV-1 transmission risk. PLoS ONE 5: e12598. doi:10.1371/journal.pone.0012598
29. Wang S, Xu F, Demirci U (2010) Advances in developing HIV-1 viral load assays for resource-limited settings. Biotechnol Adv 28: 770–781.
30. de Wolf F, Spijkerman I, Schellekens PT, Langendam M, Kuiken C, et al. (1997) AIDS prognosis based on HIV-1 RNA, CD4+ T-cell count and function: markers with reciprocal predictive value over time after seroconversion. AIDS 11: 1799–1806.
31. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250
32. Novitsky V, Wang R, Bussmann H, Lockman S, Baum M, et al. (2010) HIV-1 subtype C-infected individuals maintaining high viral load as potential targets for the "test-and-treat" approach to reduce HIV transmission. PLoS ONE 5: e10148. doi:10.1371/journal.pone.0010148
33. World Health Organization (2010) Antiretroviral drugs for treating pregnant women and preventing HIV infection in infants: recommendations for a public health approach—2010 version. Geneva: World Health Organization.
34. Schouten EJ, Jahn A, Midiani D, Makombe SD, Mnthambala A, et al. (2011) Prevention of mother-to-child transmission of HIV and the health-related Millennium Development Goals: time for a public health approach. Lancet 378: 282–284.
35. Lesotho Ministry of Health and Social Welfare, ICF Macro (2010) Lesotho Demographic and Health Survey 2009. Calverton (Maryland): Ministry of Health and Social Welfare and ICF Macro.
36. Malawi National Statistical Office, ICF Macro (2011) Malawi Demographic and Health Survey 2010. Calverton (Maryland): Malawi National Statistical Office and ICF Macro.
37. Kenya National Bureau of Statistics, ICF Macro (2010) Kenya Demographic and Health Survey 2008-09. Calverton (Maryland): Kenya National Bureau of Statistics and ICF Macro.
38. Mugo NR, Heffron R, Donnell D, Wald A, Were EO, et al. (2011) Increased risk of HIV-1 transmission in pregnancy: a prospective study among African HIV-1-serodiscordant couples. AIDS 25: 1887–1895.
39. Patel D, Cortina-Borja M, Thorne C, Newell ML (2007) Time to undetectable viral load after highly active antiretroviral therapy initiation among HIV-infected pregnant women. Clin Infect Dis 44: 1647–1656.
40. United Nations Department of Economic and Social Affairs Population Division (2011) World population prospects, the 2010 revision. New York: United Nations. Available: http://esa.un.org/unpd/wpp/index.htm. Accessed 19 December 2011.
41. Ekouevi DK, Inwoley A, Tonwe-Gold B, Danel C, Becquet R, et al. (2007) Variation of CD4 count and percentage during pregnancy and after delivery: implications for HAART initiation in resource-limited settings. AIDS Res Hum Retroviruses 23: 1469–1474.
42. Côte d'Ivoire Institut National de la Statistique, ORC Macro (2006) Enquête sur les indicateurs du sida, Côte d'Ivoire 2005. Calverton (Maryland): Côte d'Ivoire Institut National de la Statistique and ORC Macro.
43. Oladokun RE, Awolude O, Brown BJ, Adesina O, Oladokun A, et al. (2010) Service uptake and performance of the prevention of mother-to-child transmission (PMTCT) programme in Ibadan, Nigeria. Afr J Med Med Sci 39: 81–87.

44. Holmes CN, Preko PO, Bolds R, Baidoo J, Jolly PE (2008) Acceptance of voluntary counselling, testing and treatment for HIV among pregnant women in Kumasi, Ghana. Ghana Med J 42: 8–15.
45. Geddes R, Giddy J, Butler LM, Van Wyk E, Crankshaw T, et al. (2011) Dual and triple therapy to prevent mother-to-child transmission of HIV in a resource-limited setting—lessons from a South African programme. S Afr Med J 101: 651–654.
46. Torpey K, Kabaso M, Kasonde P, Dirks R, Bweupe M, et al. (2010) Increasing the uptake of prevention of mother-to-child transmission of HIV services in a resource-limited setting. BMC Health Serv Res 10: 29.
47. Bancheno WM, Mwanyumba F, Mareverwa J (2010) Outcomes and challenges of scaling up comprehensive PMTCT services in rural Swaziland, Southern Africa. AIDS Care 22: 1130–1135.
48. Hensen B, Baggaley R, Wong VJ, Grabbe KL, Shaffer N, et al. (2011) Universal voluntary HIV testing in antenatal care settings: a review of the contribution of provider-initiated testing & counselling. Trop Med Int Health. E-pub ahead of print. doi:10.1111/j.1365-3156.2011.02893.x
49. Ferguson L, Grant AD, Watson-Jones D, Kahawita T, Ong'ech JO, et al. (2012) Linking women who test HIV-positive in pregnancy-related services to long-term HIV care and treatment services: a systematic review. Trop Med Int Health. E-pub ahead of print. doi:10.1111/j.1365-3156.2012.02958.x
50. Toro PL, Katyal M, Carter RJ, Myer L, El-Sadr WM, et al. (2010) Initiation of antiretroviral therapy among pregnant women in resource-limited countries: CD4+ cell count response and program retention. AIDS 24: 515–524.
51. Boyles TH, Wilkinson LS, Leisegang R, Maartens G (2011) Factors influencing retention in care after starting antiretroviral therapy in a rural South African programme. PLoS ONE 6: e19201. doi:10.1371/journal.pone.0019201
52. Bwirire LD, Fitzgerald M, Zachariah R, Chikafa V, Massaquoi M, et al. (2008) Reasons for loss to follow-up among mothers registered in a prevention-of-mother-to-child transmission program in rural Malawi. Trans R Soc Trop Med Hyg 102: 1195–1120.
53. Johri M, Ako-Arrey D (2011) The cost-effectiveness of preventing mother-to-child transmission of HIV in low- and middle-income countries: systematic review. Cost Eff Resour Alloc 9: 3.
54. Shah M, Johns B, Abimiku A, Walker DG (2011) Cost-effectiveness of new WHO recommendations for prevention of mother-to-child transmission of HIV in a resource-limited setting. AIDS 25: 1903–1102.
55. United Nations Development Programme (2010) What will it take to achieve the Millennium Development Goals? An international assessment. New York: United Nations Development Programme.
56. van den Akker T, Bemelmans M, Ford N, Jemu M, Diggle E, et al. (2012) HIV care need not hamper maternity care: a descriptive analysis of integration of services in rural Malawi. BJOG 199: 431–438.
57. Abdool Karim SS, Naidoo K, Grobler A, Padayatchi N, Baxter C, et al. (2010) Timing of initiation of antiretroviral drugs during tuberculosis therapy. New Engl J Med 362: 697–706.
58. World Health Organization (2011) Global tuberculosis control: WHO report 2011. Geneva: World Health Organization.
59. Goletti D, Weissman D, Jackson RW, Graham NM, Vlahov D, et al. (1996) Effect of Mycobacterium tuberculosis on HIV replication. Role of immune activation. J Immunol 157: 1271–1278.
60. Toossi Z, Mayanja-Kizza H, Hirsch CS, Edmonds KL, Spahlinger T, et al. (2001) Impact of tuberculosis (TB) on HIV-1 activity in dually infected patients. Clin Exp Immunol 123: 233–238.
61. Wolday D, Hailu B, Girma M, Hailu E, Sanders E, et al. (2003) Low CD4+ T-cell count and high HIV viral load precede the development of tuberculosis disease in a cohort of HIV-positive ethiopians. Int J Tuberc Lung Dis 7: 110–116.
62. Day JH, Grant AD, Fielding KL, Morris L, Moloi V, et al. (2004) Does tuberculosis increase HIV load? J Infect Dis 190: 1677–1684.
63. Mahan CS, Walusimbi M, Johnson DF, Lancioni C, Charlebois E, et al. (2010) Tuberculosis treatment in HIV infected Ugandans with CD4 counts>350 cells/mm reduces immune activation with no effect on HIV load or CD4 count. PLoS ONE 5: e9138. doi:10.1371/journal.pone.0009138
64. Williams BG, Granich R, De Cock KM, Glaziou P, Sharma A, et al. (2010) Antiretroviral therapy for tuberculosis control in nine African countries. Proc Natl Acad Sci U S A 107: 19485–19489.
65. Badri M, Wilson D, Wood R (2002) Effect of highly active antiretroviral therapy on incidence of tuberculosis in South Africa: a cohort study. Lancet 359: 2059–2064.
66. World Health Organization (2009) Global tuberculosis control 2009: epidemiology, strategy, financing. Geneva: World Health Organization.
67. Gandhi NR, Moll AP, Lalloo U, Pawinski R, Zeller K, et al. (2009) Successful integration of tuberculosis and HIV treatment in rural South Africa: the Sizonq'oba study. J Acquir Immune Defic Syndr 50: 37–43.
68. Sterling TR, Pham PA, Chaisson RE (2010) HIV infection-related tuberculosis: clinical manifestations and treatment. Clin Infect Dis 50 (Suppl 3): S223–S230.
69. World Health Organization (2012) Guidance on couples HIV Testing and counselling including antiretroviral therapy for treatment and prevention in serodiscordant couples. Geneva: World Health Organization.
70. Ware NC (2011) What's love got to do with it?—a relationship framework for research on adherence to antiretroviral therapy (ART) as pre-exposure prophylaxis [abstract]. 6th International Conference on HIV Treatment and Prevention Adherence; 22–24 May 2011; Miami Beach, Florida.
71. Cohen MS, Muessig KE, Smith MK, Powers K, Kashuba AD (2012) Antiviral agents and HIV prevention: controversies, conflicts and consensus. AIDS. E-pub ahead of print. doi:10.1097/QAD.0b013e3283543e83
72. Hallett TB, Baeten JM, Heffron R, Barnabas R, de Bruyn G, et al. (2011) Optimal uses of antiretrovirals for prevention in HIV-1 serodiscordant heterosexual couples in South Africa: a modelling study. PLoS Med 8: e1001123. doi:10.1371/journal.pmed.1001123
73. Reynolds SJ, Makumbi F, Nakigozi G, Kagaayi J, Gray RH, et al. (2011) HIV-1 transmission among HIV-1 discordant couples before and after the introduction of antiretroviral therapy in Rakai, Uganda. AIDS 25: 473–477.
74. El-Sadr WM, Coburn BJ, Blower S (2011) Modeling the impact on the HIV epidemic of treating discordant couples with antiretrovirals to prevent transmission. AIDS 25: 2295–2299.
75. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
76. Lingappa JR, Lambdin B, Bukusi EA, Ngure K, Kavuma L, et al. (2008) Regional differences in prevalence of HIV-1 discordance in Africa and enrollment of HIV-1 discordant couples into an HIV-1 prevention trial. PLoS ONE 3: e1411. doi:10.1371/journal.pone.0001411
77. Chemaitelly H, Cremin I, Shelton J, Hallett T, Abu-Raddad LJ (2012) Distinct HIV discordancy patterns by epidemic size in stable sexual partnerships in sub-Saharan Africa. Sex Transm Infect 88: 51–57.
78. Tumwesigye E, Baeten J, Tumwebaze H, Kurth A, Revall J, et al. (2011) Potential of household-based HIV counseling and testing as a platform for targeted referral to HIV prevention and care in Uganda [abstract]. 6th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; 17–20 July 2011; Rome, Italy.
79. Joint United Nations Programme on HIV/AIDS (2011) UNAIDS terminology guidelines. Geneva: Joint United Nations Programme on HIV/AIDS.
80. Pisani E, Garnett GP, Brown T, Stover J, Grassly NC, et al. (2003) Back to basics in HIV prevention: focus on exposure. BMJ 326: 1384–1387.
81. Vickerman P, Foss AM, Pickles M, Deering K, Verma S, et al. (2010) To what extent is the HIV epidemic in southern India driven by commercial sex? A modelling analysis. AIDS 24: 2563–2572.
82. Boily MC, Lowndes C, Alary M (2002) The impact of HIV epidemic phases on the effectiveness of core group interventions: insights from mathematical models. Sex Transm Infect 78 (Suppl 1): i78–i90.
83. Cote AM, Sobela F, Dzokoto A, Nzambi K, Asamoah-Adu C, et al. (2004) Transactional sex is the driving force in the dynamics of HIV in Accra, Ghana. AIDS 18: 917–925.
84. Alary M, Lowndes CM (2004) The central role of clients of female sex workers in the dynamics of heterosexual HIV transmission in sub-Saharan Africa. AIDS 18: 945–947.
85. Gouws E, White PJ, Stover J, Brown T (2006) Short term estimates of adult HIV incidence by mode of transmission: Kenya and Thailand as examples. Sex Transm Infect 82 (Suppl 3): iii51–iii55.
86. Quan VM, Chung A, Long HT, Dondero TJ (2000) HIV in Vietnam: the evolving epidemic and the prevention response, 1996 through 1999. J Acquir Immune Defic Syndr 25: 360–369.
87. Park LS, Siraprapasiri T, Peerapatanapokin W, Manne J, Niccolai L, et al. (2010) HIV transmission rates in Thailand: evidence of HIV prevention and transmission decline. J Acquir Immune Defic Syndr 54: 430–436.
88. Lowndes CM, Alary M, Belleau M, Bosu W, Kintin D, et al. (2008) West Africa HIV/AIDS epidemiology and response synthesis. Characterisation of the HIV epidemic and response in West Africa: implications for prevention. Washington (District of Columbia): World Bank.
89. Ng M, Gakidou E, Levin-Rector A, Khera A, Murray CJ, et al. (2011) Assessment of population-level effect of Avahan, an HIV-prevention initiative in India. Lancet 378: 1643–1652.
90. Shahmanesh M, Patel V, Mabey D, Cowan F (2008) Effectiveness of interventions for the prevention of HIV and other sexually transmitted infections in female sex workers in resource poor setting: a systematic review. Trop Med Int Health 13: 659–679.
91. McClelland L, Wanje G, Kashonga F, Kibe L, McClelland RS, et al. (2011) Understanding the context of HIV risk behavior among HIV-positive and HIV-negative female sex workers and male bar clients following antiretroviral therapy rollout in Mombasa, Kenya. AIDS Educ Prev 23: 299–312.
92. Braunstein SL, Umulisa MM, Veldhuijzen NJ, Kestelyn E, Ingabire CM, et al. (2011) HIV diagnosis, linkage to HIV care, and HIV risk behaviors among newly diagnosed HIV-positive female sex workers in Kigali, Rwanda. J Acquir Immune Defic Syndr 57: e70–e76.
93. Huet C, Ouedraogo A, Konate I, Traore I, Rouet F, et al. (2011) Long-term virological, immunological and mortality outcomes in a cohort of HIV-infected female sex workers treated with highly active antiretroviral therapy in Africa. BMC Public Health 11: 700.
94. McClelland RS, Graham SM, Richardson BA, Peshu N, Masese LN, et al. (2010) Treatment with antiretroviral therapy is not associated with increased sexual risk behavior in Kenyan female sex workers. AIDS 24: 891–897.
95. McKinnon LR, Kimani M, Wachihi C, Nagelkerke NJ, Muriuki FK, et al. (2010) Effect of baseline HIV disease parameters on CD4+ T cell recovery after antiretroviral therapy initiation in Kenyan women. PLoS ONE 5: e11434. doi:10.1371/journal.pone.0011434

96. Deering KN, Shannon K, Sinclair H, Parsad D, Gilbert E, et al. (2009) Piloting a peer-driven intervention model to increase access and adherence to antiretroviral therapy and HIV care among street-entrenched HIV-positive women in Vancouver. AIDS Patient Care STDS 23: 603–609.
97. Shannon K, Bright V, Duddy J, Tyndall MW (2005) Access and utilization of HIV treatment and services among women sex workers in Vancouver's Downtown Eastside. J Urban Health 82: 488–497.
98. Chan CH, McCabe CJ, Fisman DN (2011) Core groups, antimicrobial resistance and rebound in gonorrhoea in North America. Sex Transm Infect 88: 200–204.
99. Kennedy C, O'Reilly K, Medley A, Sweat M (2007) The impact of HIV treatment on risk behaviour in developing countries: a systematic review. AIDS Care 19: 707–720.
100. Montague BT, Vuylsteke B, Buve A (2011) Sustainability of programs to reach high risk and marginalized populations living with HIV in resource limited settings: implications for HIV treatment and prevention. BMC Public Health 11: 701.
101. Verma R, Shekhar A, Khobragade S, Adhikary R, George B, et al. (2010) Scale-up and coverage of Avahan: a large-scale HIV-prevention programme among female sex workers and men who have sex with men in four Indian states. Sex Transm Infect 86 (Suppl 1): i76–i82.
102. Alary M, Mukenge Tshibaka L, Bernier F, Geraldo N, Lowndes CM, et al. (2002) Decline in the prevalence of HIV and sexually transmitted diseases among female sex workers in Cotonou, Benin, 1993–1999. AIDS 16: 463–470.
103. Ghys PD, Diallo MO, Ettiegne Traore V, Kale K, Tawil O, et al. (2002) Increase in condom use and decline in HIV and sexually transmitted diseases among female sex workers in Abidjan, Cote d'Ivoire, 1991–1998. AIDS 16: 251–258.
104. Chandrashekar S, Guinness L, Kumaranayake L, Reddy B, Govindraj Y, et al. (2010) The effects of scale on the costs of targeted HIV prevention interventions among female and male sex workers, men who have sex with men and transgenders in India. Sex Transm Infect 86 (Suppl 1): i89–i94.
105. Creese A, Floyd K, Alban A, Guinness L (2002) Cost-effectiveness of HIV/AIDS interventions in Africa: a systematic review of the evidence. Lancet 359: 1635–1643.
106. Vickerman P, Terris-Prestholt F, Delany S, Kumaranayake L, Rees H, et al. (2006) Are targeted HIV prevention activities cost-effective in high prevalence settings? Results from a sexually transmitted infection treatment project for sex workers in Johannesburg, South Africa. Sex Transm Dis 33 (Suppl 10): S122–S132.
107. Diabate S, Zannou D, Geraldo N, Chamberland A, Akakpo J, et al. (2011) Antiretroviral therapy among HIV-1 infected female sex workers in Benin: a comparative study with patients from the general population. World J AIDS 1: 94–99.
108. Rodriguez-Arenas MA, Jarrin I, del Amo J, Iribarren JA, Moreno S, et al. (2006) Delay in the initiation of HAART, poorer virological response, and higher mortality among HIV-infected injecting drug users in Spain. AIDS Res Hum Retroviruses 22: 715–723.
109. Watts C, Zimmerman C, Foss AM, Hossain M, Cox A, et al. (2010) Remodelling core group theory: the role of sustaining populations in HIV transmission. Sex Transm Infect 86 (Suppl 3): iii85–iii92.
110. Boily MC, William J, Alary M (2011) Impact of female sex worker turnover on HIV prevalence, incidence and core group intervention impact in Cotonou (Benin) [poster]. 19th Biennial Conference of the International Society for Sexually Transmitted Diseases Research; 10–14 July 2011; Quebec City, Canada.
111. Pickles M, Boily MC, Vickerman P, Ramesh BM, Washington R, et al. (2011) Time evolution of the fraction of new HIV infections due to primary infection among high risk groups in southern India [poster]. 19th Biennial Conference of the International Society for Sexually Transmitted Diseases Research; 10–14 July 2011; Quebec City, Canada.
112. Vickerman P, Watts C, Delany S, Alary M, Rees H, et al. (2006) The importance of context: model projections on how microbicide impact could be affected by the underlying epidemiologic and behavioral situation in 2 African settings. Sex Transm Dis 33: 397–405.
113. Boily MC, Lowndes C, Alary M (2000) Complementary hypothesis concerning the community sexually transmitted disease mass treatment puzzle in Rakai, Uganda. AIDS 14: 2583–2592.
114. Geibel S, van der Elst EM, King'ola N, Luchters S, Davies A, et al. (2007) 'Are you on the market?': a capture-recapture enumeration of men who sell sex to men in and around Mombasa, Kenya. AIDS 21: 1349–1354.
115. van Griensven F (2007) Men who have sex with men and their HIV epidemics in Africa. AIDS 21: 1361–1362.
116. Joint United Nations Programme on HIV/AIDS (2010) New HIV infections by mode of transmission in West Africa: A multicountry analysis. Geneva: Joint United Nations Programme on HIV/AIDS.
117. Mathers B, Degenhardt L, Phillips B, Wiessing L, Hickman M, et al. (2008) Global epidemiology of injecting drug use and HIV among people who inject drugs: a systematic review. Lancet 372: 1733–1745.
118. Strathdee SA, Hallett TB, Bobrova N, Rhodes T, Booth R, et al. (2010) HIV and risk environment for injecting drug users: the past, present, and future. Lancet 376: 268–284.
119. Joint United Nations Programme on HIV/AIDS (2009) KENYA HIV prevention response and modes of transmission analysis. Geneva: Joint United Nations Programme on HIV/AIDS.
120. Wolfe D, Carrieri MP, Shepard D (2010) Treatment and care for injecting drug users with HIV infection: a review of barriers and ways forward. Lancet 376: 355–366.
121. Malta M, Magnanini MM, Strathdee SA, Bastos FI (2010) Adherence to antiretroviral therapy among HIV-infected drug users: a meta-analysis. AIDS Behav 14: 731–747.
122. Ndiaye B, Ould-Kaci K, Salleron J, Bataille P, Bonnevie F, et al. (2009) Incidence rate and risk factors for loss to follow-up in HIV-infected patients from five French clinical centres in Northern France—January 1997 to December 2006. Antivir Ther 14: 567–575.
123. Mocroft A, Kirk O, Aldins P, Chies A, Blaxhult A, et al. (2008) Loss to follow-up in an international, multicentre observational study. HIV Med 9: 261–269.
124. Rosen S, Sanne I, Collier A, Simon JL (2005) Rationing antiretroviral therapy for HIV/AIDS in Africa: choices and consequences. PLoS Med 2: e303. doi:10.1371/journal.pmed.0020303
125. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469.
126. International Maternal Pediatric Adolescent AIDS Clinical Trials Group (2010) PROMISE featured study—1077HS. Available: https://impaactgroup.org/promise-featured-study-1077hs. Accessed June 8, 2012.
127. Cornell M, Grimsrud A, Fairall L, Fox MP, van Cutsem G, et al. (2010) Temporal changes in programme outcomes among adult patients initiating antiretroviral therapy across South Africa, 2002–2007. AIDS 24: 2263–2270.
128. Nachega JB, Morroni C, Zuniga JM, Schechter M, Rockstroh J, et al. (2012) HIV treatment adherence, patient health literacy, and health care provider-patient communication: results from the 2010 AIDS Treatment for Life International Survey. J Int Assoc Physicians AIDS Care (Chic) 11: 128–133.
129. Thompson MA, Mugavero MJ, Amico KR, Chang LW, Gross R, et al. (2012) Guidelines for improving entry into and retention in care and antiretroviral adherence for persons with HIV: evidence-based recommendations from an International Association of Physicians in AIDS Care panel. Ann Intern Med. Ann Intern Med 156: 817–833.
130. World Health Organization (2003) Making choices in health: generalized cost effectiveness analysis: a guide. Geneva: World Health Organization.
131. World Health Organization (2011) WHO consultation on the strategic use of antiretrovirals for treatment and prevention of HIV. Geneva: World Health Organization.
132. Lewden C, Gabillard D, Minga A, Ekouevi DK, Avit D, et al. (2012) CD4-specific mortality rates among HIV-infected adults with high CD4 counts and no antiretroviral treatment in West Africa. J Acquir Immune Defic Syndr 59: 213–219.

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Debate and Commentary—Will Early Infection Compromise Treatment-as-Prevention Strategies?

Myron S. Cohen[1,2,3¶], Christopher Dye[4¶], Christophe Fraser[5¶*], William C. Miller[2,3¶], Kimberly A. Powers[2,3¶*], Brian G. Williams[6¶*]

1 Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina, United States of America, **2** Department of Medicine, University of North Carolina, Chapel Hill, North Carolina, United States of America, **3** Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, United States of America, **4** World Health Organization, Geneva, Switzerland, **5** Medical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **6** South African Centre for Epidemiological Modelling and Analysis, Stellenbosch, South Africa

Abstract: Universal HIV testing and immediate antiretroviral therapy for infected individuals has been proposed as a way of reducing the transmission of HIV and thereby bringing the HIV epidemic under control. It is unclear whether transmission during early HIV infection—before individuals are likely to have been diagnosed with HIV and started on antiretroviral therapy—will compromise the effectiveness of treatment as prevention. This article presents two opposing viewpoints by Powers, Miller, and Cohen, and Williams and Dye, followed by a commentary by Fraser.

Introduction to the Debate

Triple-combination antiretroviral therapy (ART) first became available in 1995 for the treatment of people living with HIV [1]. The fact that ART reduces viral load raised the prospect of using ART not only to keep people alive, but also to control the epidemic by annually testing everyone at risk of HIV infection and immediately starting infected individuals on ART [2–4]. In 2011, the HPTN 052 trial showed that ART reduced the risk of infection in heterosexual HIV-serodiscordant couples (where one individual is HIV-seropositive and the other is not) by 96% (95% confidence interval [CI], 73% to 99%) and decisively confirmed the impact of treatment on heterosexual transmission [5].

If the individual-level effects observed in the HPTN 052 trial are to be successfully replicated at a population level, many operational issues need to be addressed. An issue of particular importance concerns transmission in the early stages of HIV infection, before an individual is likely to be diagnosed with HIV and start receiving treatment. Early HIV infection (EHI), the first 3–6 months after infection, includes acute HIV infection (AHI), the period before the development of antibodies to HIV, when the concentration of virus in the plasma spikes and then falls to the set-point viral load. Chronic HIV infection (CHI) comprises an asymptomatic period following EHI, characterized by a stable viral load ($\sim 10^{4.5}$ copies/ml) and relatively low transmissibility, as well as late infection/AIDS, when viral load and transmissibility are elevated again. Different studies have arrived at widely differing estimates of the proportion of HIV transmission events that occur during the first 3–6 months after HIV infection, ranging from 5% to 95% [6]. High levels of HIV transmission early in infection could compromise the impact of universal testing and treatment on HIV transmission at a population level, so it is essential to resolve this issue if ART is to be used to help control the HIV pandemic.

In this debate, which specifically addresses heterosexual transmission of HIV, Powers, Miller, and Cohen argue that up to 40% of transmission takes place during EHI, and that this transmission will compromise the effectiveness of treatment as prevention. Williams and Dye argue that only about 2% of transmission takes place during AHI, so that annual testing and immediate ART will be sufficient to control the epidemic of HIV. In the final section, Christophe Fraser summarizes and weighs in on the debate.

Kimberly A. Powers, William C. Miller, Myron S. Cohen's Viewpoint: Acute and Early HIV Infection Will Limit the Effectiveness of HIV Treatment as Prevention

Although the use of ART to stop the spread of HIV has become a major focus of HIV prevention, reliable empirical evidence to support this strategy at the population level does not exist, and its success in the real world may be limited by numerous factors [7].

Citation: Cohen MS, Dye C, Fraser C, Miller WC, Powers KA, et al. (2012) HIV Treatment as Prevention: Debate and Commentary—Will Early Infection Compromise Treatment-as-Prevention Strategies? PLoS Med 9(7): e1001232. doi:10.1371/journal.pmed.1001232

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: WCM, MSC and KAP received financial support from the National Institutes of Health (NIH) through grant R01 AI083059. CF received financial support from the Royal Society and the Medical Research Council. MSC received financial support from the HIV Prevention Trials Network (HPTN 052), the National Institutes of Health (grant number R37 DK049381), the Center for HIV/AIDS Vaccine Immunology, and the University of North Carolina at Chapel Hill Center for AIDS Research (CFAR), and an NIH-funded program (grant number P30 AI50410). CD is a staff member of the World Health Organization. The authors alone are responsible for the views expressed in this publication, which do not necessarily represent the decisions, policy or views of the World Health Organization. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: c.fraser@imperial.ac.uk (CF); powersk@email.unc.edu (KAP); BrianGerard Williams@gmail.com (BGW)

¶ All authors contributed equally to this work.

Abbreviations: AHI, acute HIV infection; ART, antiretroviral therapy; CHI, chronic HIV infection; CI, confidence interval; EHI, early HIV infection

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Here we focus on one particular factor: transmission during AHI and EHI, which will not be affected by a treatment-as-prevention approach. We argue that high levels of transmission during this period of maximal infectiousness [6] will compromise the effectiveness of HIV treatment as prevention.

Transmission Biology

HIV acquisition leads to a ramp-up in viremia to 10 million copies/ml or more [8], with a subsequent cell-mediated immune response that leads to decreased viral replication during asymptomatic infection [9]. The best available estimates of heterosexual HIV transmission by infection stage, calculated among steady couples in Rakai, Uganda, suggest that transmissibility is 26 times as high (95% CI, 13–54) during EHI as it is in the subsequent asymptomatic period [10]. Among the Rakai couples, the probability of a newly infected person transmitting HIV to his or her partner in the first five months of infection was estimated at 43% [11].

Importantly, the elevation in transmissibility observed during EHI in Rakai is greater than would be expected based on viral load alone [10]. If viral load were the only driver of infectiousness, then we would expect transmission rates during AHI and EHI to be only a few times higher than during chronic infection, as Williams and Dye describe below. The mechanism for the additional enhancement in transmissibility observed during EHI in Rakai has not been elucidated, but there is evidence from macaques that individual virions from EHI are 75–750 times as infectious as virions from CHI [12].

Mathematical Modelling

Mathematical modelers have attempted to predict the potential population benefits of ART [13]. With perhaps the most optimistic model, Granich et al. have argued that universal annual HIV testing and immediate ART would lead to HIV "elimination," defined as one incident infection per 1,000 persons annually, within ten years in South Africa [4]. However, the analyses leading to this conclusion failed to account for the effect of poor engagement in care [14] and the increased infectiousness of persons with EHI [15], who would not be reached by the test-and-treat strategy.

Modelling estimates of the percentage of new cases that are due to contact with EHI index cases vary widely, depending on epidemic stage, model structure, transmission mode, and EHI definition. Most endemic-phase estimates have been in the range of 5% to 40% [6], broadly consistent with estimates of 25%–50% from phylogenetic studies [16]. However, the data available for parameterizing most of these models have been limited. Using behavioral and viral load data from Lilongwe, Malawi, as well as the best available estimates of transmission efficiency by infection stage [10], we constructed a mathematical model that allowed for transmission both within and outside of steady heterosexual relationships, substantial variation in transmissibility over the course of infection, and heterogeneity in behavioral risk (rates of partner change and contact within pairs) [15]. We used a Bayesian melding procedure to account for input uncertainty, to fit the model to empirical HIV prevalence data from Lilongwe, and to express uncertainty around outputs.

In our model, we estimated that transmission rates were elevated for 4.8 months (i.e., EHI lasted for 4.8 months) and that 38% (95% credible interval 19%–52%) of endemic-phase incident infections arose from contact with EHI index cases annually. We estimated that an annual test-and-treat strategy with very optimistic chronic-phase coverage and engagement levels of 90% or greater could substantially reduce HIV incidence, and eventually prevalence, in this setting, but HIV elimination was possible only at coverage and engagement levels of 99% or greater. This prediction is consistent with the conclusion of Granich et al. [4] and Williams and Dye (below) that with truly universal coverage and engagement, annual test-and-treat strategies could lead to HIV elimination. At more realistic coverage and engagement levels (75%–85% of chronic-phase cases), however, elimination did not occur in our model, but additional interventions halting transmission during EHI led to marked, durable reductions in HIV prevalence and incidence. Even in sensitivity analyses where the contribution of EHI in Lilongwe was only half our best estimate of 38% (19%), our results suggested that unless test-and-treat coverage is essentially perfect, the impact of such interventions is likely to be limited substantially by transmission during EHI.

Williams et al. [17] recently argued (as Williams and Dye argue below) that the 38% of cases we estimated to arise from contact with EHI index cases is too high, proposing instead that only 2%–4% of incident infections arise during AHI. By basing their calculations only on the putative relationship between chronic-phase viral load and transmissibility, they do not capture the greater-than-expected transmissibility observed during EHI among the Rakai couples [10]. Furthermore, the duration (one month) and increase in transmission rate per sexual encounter (three-fold) that they calculate for AHI correspond to an expected within-couple transmission probability of $\sim$3% during AHI (calculated as $1-e^{-\beta\tau d}$, where $\beta = 0.106$ cases per person-year, the asymptomatic-period transmission rate estimated from the Rakai data [10]; $\tau = 3$, the proposed relative increase in transmissibility comparing AHI and asymptomatic infection [17] and below; and $d = 1$ month, the proposed duration of AHI [17] and below). This within-couple transmission probability of 3% during AHI is dramatically lower than the 43% observed during EHI in Rakai [11]. Simply put, the calculations of Williams and Dye are inconsistent with the best available data from epidemiological, mathematical, and phylogenetic studies regarding transmission during EHI [6,10,11,16].

Implications for Treatment as Prevention

We believe that EHI can be expected to limit the impact of treatment-as-prevention programs—at least in settings similar to Lilongwe, Malawi—and that reductions in HIV incidence and prevalence can be optimized through intervention packages that stop transmission during both CHI and EHI. A number of randomized trials to investigate the population-level effects of treatment as prevention are underway [18]. Because they do not include a specific strategy for dealing with AHI or EHI, the extent to which they succeed will provide some indication as to whether or not transmission during EHI compromises the effectiveness of treatment as prevention. In addition, some of these studies will use phylogenetic measurements to clarify transmission events attributable to acute/early cases versus chronic cases, providing more specific information about the importance of AHI/EHI in the context of these trials. If AHI and EHI are found to limit treatment as prevention empirically, we will need to develop a more efficient strategy for identifying individuals with EHI, as well as credible behavioral and/or treatment-based intervention strategies [19] for this period.

Brian G. Williams and Christopher Dye's Viewpoint: Acute and Early HIV Infection Will Not Limit the Effectiveness of HIV Treatment as Prevention

It has been shown that successful ART reduces heterosexual transmission of HIV by 96% (95% CI, 73%–99%) [5], more than enough to eliminate HIV transmission [4]. However, Powers et al.

([15] and above) argue that 38% of transmission events occur during the first 4.8 months after HIV acquisition and that unless HIV-positive individuals start ART very early in this period, ART will not be sufficient to eliminate HIV transmission. Here we argue that (1) there is no convincing evidence to support their estimate of the proportion of transmission events that take place during the first 4.8 months, and (2) even if their estimate is correct, annual testing and immediate treatment would nevertheless be sufficient to eliminate transmission

Transmission during Early HIV Infection

Twelve studies, summarized by Cohen et al. [6], suggest that between 8% and 75% of new infections occur during EHI. Unfortunately, all of the studies that are concerned with heterosexual transmission depend on one set of data collected from the retrospective identification of 23 couples in Rakai, Uganda [11], in a study designed for other purposes. In ten couples in the Rakai study, both partners seroconverted in the same ten-month interval between testing. It was assumed that the first person in the couple was infected after an average of five months, leaving five months for them to infect their partner. Of the 13 remaining serodiscordant couples, three of the seronegative partners were infected in the next ten months, giving a rate ratio for infection during the first and second periods of 7.3 (95% CI, 3.1–17.3) [11]. Allowance was made for the self-reported number of sexual encounters, but not for the possibility that the second person was infected from outside the relationship. Data from a study of 23 couples, designed for other purposes, in which people were tested for HIV only at ten-month intervals and were identified retrospectively, which relied on self-reported sexual activity, and which did not determine whether or not the infection came from outside the relationship [11], do not provide a sound basis for drawing conclusions about the importance of EHI.

Since there is no convincing direct evidence that heterosexual transmission is higher during AHI than during the asymptomatic period of CHI, we consider indirect estimates based on viral load and the likely duration of AHI. Most new HIV infections are established by a single founder virus. The concentration of virions in the plasma then increases rapidly over three to four weeks, reaching $\sim 10^{6.5}$ copies/ml, and then falls equally rapidly to a set point at $\sim 10^{4.5}$ copies/ml [20,21]. From a preliminary analysis of data presented by Robb [20] for people in the acute phase of infection, the peak concentration of virus in the plasma is $10^{2.1}$ (95% CI,$10^{1.6}$–$10^{2.5}$) copies/ml times greater than at the set point, and AHI lasts for 2.1 (1.6–2.5) weeks. Miller et al. [16] likewise observe that "acute HIV infection, when the concentration of HIV in blood and genital secretions is extremely high, is only a few weeks in duration." According to the model of Powers et al. (Supplementary Web Appendix Figure 1 in [15]), AHI , the period of peak viral load that lasts for a maximum of six weeks, corresponding to an average duration, with the same area under the curve, of two weeks, with average viral load increased about 20-fold.

Transmission increases with viral load, and most authors assume that transmission increases as viral load to the power of 0.3 to 0.5 [22–24]; the relationship is clearly sublinear so that transmission saturates as viral load increases [25]. A more biologically plausible model [26], which gives an equally good fit to the available data [27], assumes that transmission increases linearly with viral load at low values of viral load, but converges to an asymptote above a viral load of $10^{4.4}$ copies/ml [27]. In order to estimate P_i, the proportion of infections that take place in stage i, we calculate, to first order,

$$P_i = \frac{r_i d_i}{\sum_i r_i d_i} \qquad (1)$$

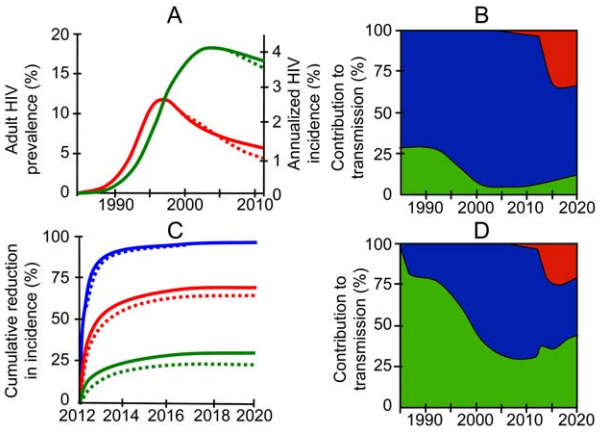

Figure 1. The predicted effect of different levels of acute infection on a combination prevention package including universal testing and treatment, as will be tested in the PopART trial [20]. (A) Green line: prevalence; red line: incidence. Two versions of a model are fitted to the adult HIV prevalence curve for South Africa (Joint United Nations Programme on HIV/AIDS): one "corrected" for serial monogamy effects in low-risk individuals [5], and thus with a low contribution of AHI (solid line), and one without the correction, and thus with a high contribution of AHI (dashed line). Fitted parameters are as follows: the proportion of individuals in three risk groups (low, medium, and high), rate of partner change for high-risk individuals, assortativity of mixing by risk, start time, early treatment rates, and an overall infectiousness parameter. Other parameters were fixed from the literature [13,14]. (C) The intervention is introduced in 2012, and predictions are made until 2020, for three scenarios ranging from the very pessimistic (green line), through "just on target" (red line), to very optimistic (blue line). The results are surprisingly independent of the amount of transmission from AHI, as indicated by the solid versus dashed lines. (B and D) The contribution to transmission from individuals in different disease stages in the just-on-target scenarios is plotted in (B), corresponding to solid lines in (A) and (C) (corrected for serial monogamy effects), and (D), corresponding to dashed lines in (A) and (C) (not corrected for serial monogamy effects). Shown are all new infections of index cases in AHI and EHI (green), of index cases in untreated CHI (blue), and of index cases in treated CHI (red), as a proportion of total new infections.
doi:10.1371/journal.pmed.1001232.g001

where d_i is the duration and r_i is the relative infectiousness of stage i, assuming a steady state, in which prevalence, incidence, and mortality are constant, and random mixing. With a mean set-point viral load of $10^{4.5}$ copies/ml, an increase in the average viremia from $10^{4.5}$ copies/ml to $10^{6.5}$ copies/ml during AHI would make little difference to the overall rate of transmission. Even if we generously assume that the viral load peak during AHI lasts for one month and that transmission rates per sexual encounter are increased three-fold during AHI, Equation 1 shows that AHI accounts for only 2% of all transmissions and would be consequential only if people had several partners in the two-week period of AHI, which is not supported by data. Raised viremia during AHI does not support the claim that EHI contributes significantly to heterosexual HIV transmission.

Powers et al. [15] estimate that EHI lasts for 4.8 months and that during this time the risk of infection per sexual encounter is 30.3 (13.6–47.1) times greater than it is during the asymptomatic period of CHI. If we grant these assumptions, Equation 1, which assumes a steady state, shows that about 56% of infections would then occur during EHI, in agreement with their estimate of 78% (95% credible interval, 68%–85%) in 1975, falling to 38% (19%–52%) in 2010. The agreement between this estimate using Equation 1 and the estimate of Powers et al. [15] shows that our

different conclusions arise from our different estimates of the duration of elevated infectiousness and transmission rates during that period, and is not due to other structural details of the model.

Early Treatment and R_0

Even if the modeled outcomes of Powers et al. [15] are correct, annual testing and immediate treatment would still be sufficient to eliminate transmission. The initial doubling time of the prevalence of HIV in the Malawi study was 1.3 years ([15]), and the greater the relative risk of transmission in EHI, the smaller must be the value of the basic reproduction number, R_0, to maintain the same initial doubling time, as follows directly from the Euler-Lotka equation [28,29]. If we suppose that transmission per sexual contact is 30 times higher during EHI than during the next ten years of CHI, as proposed by Powers et al. [15], the value of R_0, subject to the constraint that the initial doubling time is 1.3 years, would have to be ~2 rather than ~5–10 [27]. Testing people at regular intervals of one year and starting them immediately on ART would reduce R_0 to 0.8 [27]; testing people more frequently would reduce it further. Thus, early treatment could still lead to elimination of HIV transmission, and adding other interventions, such as male circumcision, would increase the impact further.

Implications for Treatment as Prevention

There is agreement that ART reduces the rate of transmission by about 25 times [5] and that this reduction is much greater than has been demonstrated with any other currently available intervention. It is unlikely that AHI or EHI significantly compromises the impact of treatment on transmission. We agree that if the intention is to start people on ART as soon as possible after they become infected with HIV, ways of detecting people in the acute phase of HIV infection would increase the impact of treatment as prevention. Whatever may be the precise details of transmission during AHI, treatment as prevention must now be the cornerstone of HIV prevention programs. Going beyond mathematical modelling, the magnitude of the effects of treatment as prevention are being evaluated in a number of field trials [18]. We expect the results of these trials to offer, for the first time, the prospect of an AIDS-free generation [30,31].

Christophe Fraser's Commentary on the Debate

The role of AHI and EHI in transmission has been debated since the early days of the HIV epidemic [32] and for much the same reason is still debated today: it seems self-evidently important but is hard to pin down. It is the subject of renewed attention in light of growing interest in treatment as prevention, because unless diagnosis can be made during AHI, most individuals will have passed through EHI before universal testing and treatment would start them on ART.

Powers et al. argue that EHI is a major driver of the epidemic, while Williams and Dye suggest a minimal role for EHI; other studies provide estimates across this range [6]. The debating parties agree that data from the Rakai study in Uganda indicate very high onward transmission in EHI, with 43% of couples found to be mutually infected at the first follow-up after neither of them was [11], and they agree that this is not consistent with expectations from viral load alone [10]. Powers et al. support the epidemiological observation (high transmission) and argue that there is no reason to believe that viral load is a good marker for infectiousness in EHI (true), while Williams and Dye support arguments based on viral load and argue that, with only 23 couples, the Rakai study [11] is, in this context, too small to draw such a strong conclusion (also true).

Key Points

- Two opposing model-based viewpoints are presented about whether transmission during early HIV infection is likely to compromise the effectiveness of treatment as prevention, i.e., using universal HIV testing and immediate ART to halt the transmission of HIV in a population.
- Powers, Miller, and Cohen's model suggests that 38% of transmission takes place in the first few months after HIV infection, i.e., before infections would be detected and treated via annual testing and treatment, making early HIV transmission a serious impediment to treatment as prevention.
- Williams and Dye argue that their model shows that the high levels of viremia during the acute phase of HIV infection do not significantly increase HIV transmission and that the risk of infection is not significantly higher during early infection than it is during chronic infection. They argue further that even if there were much higher rates of transmission in the acute and early stages of infection, early treatment would still be effective in controlling the epidemic of HIV.
- Fraser highlights that the epidemiological contribution of acute infection depends not just on infectiousness but also on patterns of risk behavior. However, Fraser largely concurs with Williams and Dye that the effect of acute and early infection on the predicted impact of universal testing and treatment may be much smaller than expected.
- All authors agree that future modelling and empirical studies will be useful in elucidating the impact of treatment as prevention on the epidemic of HIV.

A pivotal point that neither party delves deeply enough into is the effect of patterns of risk behavior. In a reanalysis of the Rakai data, Hollingsworth et al. [10] show that low-risk (serial monogamy) and high-risk (random mixing) contexts led to significantly different estimates of the extent of transmission during EHI; Eaton et al. [33] show that transmission in a sexual network with concurrent partnerships produces intermediate estimates. Powers et al. obtained higher estimates by allowing for complex correlations between partner change rates and transmission probabilities per sex act [15]. Finally, Koopman et al. [34] emphasize that assuming constant sexual risk behavior over individuals' lifetimes is neither sensible nor supported by the data, and this too plays into the estimation of the role of AHI, since if partner change rates decline with age, EHI becomes more important.

While the role of different patterns of risk behavior in driving EHI may have been underestimated, the argument made by Williams and Dye that AHI and EHI do not matter to prevention efforts as much as we might think may in fact be more fundamental. This argument is based on the Euler-Lotka equation, which constrains the relationship between growth rates and generation times [35,36]. Here, I test this argument using a conventional mathematical model of HIV transmission, which extends earlier models [4,37] and is more complex than the models that Williams and Dye have used in this context. Estimates of transmission rates during EHI in the model are based on the data from Rakai, which is still the best evidence to date on this topic, and the model is fit to national surveillance data from South Africa (from the Joint United Nations Programme on HIV/AIDS). The model allows the rate of transmission during EHI to be modified by turning on or off the correction factor for finite

partnerships amongst low- and medium-risk individuals [10]; this is more efficient than increasing the parameter for the infectiousness of EHI, due to the counteracting effects of limited partnership turnover on biological infectiousness. When the contribution of EHI is tuned up or down, very different model projections result, as expected.

However, changing our assumptions about the importance of AHI and EHI not only affects our predictions about the future, but also changes our interpretation of what has happened in the past: each time the contribution of EHI is tuned up or down, the model must be refitted to data. Figure 1 shows the outcome of this process: it broadly confirms the prediction of the Euler-Lotka equation in the context of a more complex mathematical model, validating the hypothesis of Williams and Dye that the total effectiveness of treatment as prevention depends surprisingly little on the effect of EHI on transmission.

It must be stated that these predictions are based on a model that is still relatively simple, and reality may yet surprise us. Further modelling work could play a useful role by determining more systematically under which circumstances the prediction of the Euler-Lotka equation is or is not expected to hold, and guiding the collection of appropriate data. Treatment as prevention holds extraordinary promise, but will also be expensive and challenging to deliver in many settings. Arguments about potential barriers to success, such as presented in this debate, need careful consideration. Population-based trials, such as PopART (HPTN 071) [38] and others [18] that are being planned, as well as more observational data, will provide much needed empirical tests of the proposal that treatment as prevention is feasible and effective.

Author Contributions

Analyzed the data: MC CD CF WCM KAP BGW. Wrote the first draft of the manuscript: MC CD CF WCM KAP BGW. Contributed to the writing of the manuscript: MC CD CF WCM KAP BGW. ICMJE criteria for authorship read and met: MC CD CF WCM KAP BGW. Agree with manuscript results and conclusions: MC CD CF WCM KAP BGW. Combined individual written contributions: BGW.

References

1. Ho DD (1995) Time to hit HIV, early and hard. N Engl J Med 333: 450–451.
2. Hosseinipour M, Cohen MS, Vernazza PL, Kashuba AD (2002) Can antiretroviral therapy be used to prevent sexual transmission of human immunodeficiency virus type 1? Clin Infect Dis 34: 1391–1395.
3. Montaner JS, Hogg R, Wood E, Kerr T, Tyndall M, et al. (2006) The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 368: 531–536.
4. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
5. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
6. Cohen MS, Shaw GM, McMichael AJ, Haynes BF (2011) Acute HIV-1 infection. N Engl J Med 364: 1943–1954.
7. Smith K, Powers KA, Kashuba AD, Cohen MS (2011) HIV-1 treatment as prevention: the good, the bad, and the challenges. Curr Opin HIV AIDS 6: 315–325.
8. Piatak M Jr, Saag MS, Yang LC, Clark SJ, Kappes JC, et al. (1993) High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science 259: 1749–1754.
9. Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, et al. (1994) Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J Virol 68: 4650–4655.
10. Hollingsworth TD, Anderson RM, Fraser C (2008) HIV-1 transmission, by stage of infection. J Infect Dis 198: 687–693.
11. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, et al. (2005) Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. J Infect Dis 191: 1403–1409.
12. Ma ZM, Stone M, Piatak M Jr, Schweighardt B, Haigwood NL, et al. (2009) High specific infectivity of plasma virus from the pre-ramp-up and ramp-up stages of acute simian immunodeficiency virus infection. J Virol 83: 3288–3297.
13. Cohen MS, Gay CL (2010) Treatment to prevent transmission of HIV-1. Clin Infect Dis 50 (Suppl 3): S85–S95.
14. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ (2011) The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis 52: 793–800.
15. Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, et al. (2011) The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet 378: 256–268.
16. Miller WC, Rosenberg NE, Rutstein SE, Powers KA (2010) Role of acute and early HIV infection in the sexual transmission of HIV. Curr Opin HIV AIDS 5: 277–282.
17. Williams BG, Granich R, Dye C (2011) Role of acute infection in HIV transmission. Lancet 378: 1913; author reply 1914–1915.
18. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Current HIV Res 9: 446–469.
19. Pettifor A, MacPhail C, Corneli A, Sibeko J, Kamanga G, et al. (2011) Continued high risk sexual behavior following diagnosis with acute HIV infection in South Africa and Malawi: implications for prevention. AIDS Behav 15: 1243–1250.
20. Robb M (2011) The Early Capture HIV Cohort Study (ECHO): prospective identification of acute HIV infection prior to peak viremia among high-risk populations [poster]. HIV Evolution, Genomics and Pathogenesis Symposium; 20–25 March 2011; Whistler, British Columbia, Canada.
21. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, et al. (2003) Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 17: 1871–1879.
22. Attia S, Egger M, Muller M, Zwahlen M, Low N (2009) Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. AIDS 23: 1397–1404.
23. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, et al. (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet 375: 2092–2098.
24. Lingappa JR, Hughes JP, Wang RS, Baeten JM, Celum C, et al. (2010) Estimating the impact of plasma HIV-1 RNA reductions on heterosexual HIV-1 transmission risk. PLoS ONE 5: e12598. doi:10.1371/journal.pone.0012598
25. Williams B, Lima V, Gouws E (2011) Modelling the impact of antiretroviral therapy on the epidemic of HIV. Curr HIV Res 9: 367–382
26. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP (2007) Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. Proc Natl Acad Sci U S A 104: 17441–17446.
27. Williams BG (2011) How important is the acute phase in HIV epidemiology? arXiv. Available: http://arxiv.org/pdf/1105.2767v1.pdf. Accessed 18 May 2012.
28. Williams BG, Dransfield RD, Brightwell R (1990) Tsetse-fly (Diptera, Glossinidae) population-dynamics and the estimation of mortality-rates from life-table data. Bull Entomol Res 80: 479–485.
29. Lotka AJ (1907) Relation between birth rates and death rates. Science 26: 21–22.
30. Goosby E (2011) An inspiring vision: creating an AIDS-free generation. The Huffington Post. 9 November 2011.
31. Clinton HR (2011 Nov 8) Remarks on 'creating an AIDS-free generation'. Washington (District of Columbia): US Department of State.
32. Jacquez JA, Koopman JS, Simon CP, Longini IM Jr (1994) Role of the primary infection in epidemics of HIV infection in gay cohorts. J Acquire Immune Defic Syndr 7: 1169–1184.
33. Eaton JW, Hallett TB, Garnett GP (2011) Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS Behav 15: 687–692.
34. Koopman JS, Jacquez JA, Welch GW, Simon CP, Foxman B, et al. (1997) The role of early HIV infection in the spread of HIV through populations. J Acquir Immune Defic Syndr Hum Retrovirol 14: 249–258.
35. Wallinga J, Lipsitch M (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci 274: 599–604.
36. Grassly NC, Fraser C (2008) Mathematical models of infectious disease transmission. Nat Rev Microbiol 6: 477–487.
37. Hallett TB, Singh K, Smith JA, White RG, Abu-Raddad LJ, et al. (2008) Understanding the impact of male circumcision interventions on the spread of HIV in southern Africa. PLoS ONE 3: e2212. doi:10.1371/journal.pone.0002212
38. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: The Utility and Limitations of Ecological Observation

M. Kumi Smith[1], Kimberly A. Powers[1,2], Kathryn E. Muessig[2], William C. Miller[1,2], Myron S. Cohen[1,2,3]*

1 Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina, United States of America, **2** Department of Medicine, University of North Carolina, Chapel Hill, North Carolina, United States of America, **3** Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina, United States of America

Abstract: Results from several observational studies of HIV-discordant couples and a randomized controlled trial (HIV Prevention Trials Network 052) show that antiretroviral therapy (ART) can greatly reduce heterosexual HIV transmission in stable HIV-discordant couples. However, such data do not prove that ART will reduce HIV incidence at the population level. Observational investigations using ecological measures have been used to support the implementation of HIV treatment for the specific purpose of preventing transmission at the population level. Many of these studies note ecological associations between measures of increased ART uptake and decreased HIV transmission. Given the urgency of implementing HIV prevention measures, ecological studies must de facto be used to inform current strategies. However, the hypothesis that widespread ART can eliminate HIV infection may have raised expectations beyond what we may be able to achieve. Here we review and discuss the construct of the exposure and outcome measures and analysis methods used in ecological studies. By examining the strengths and weaknesses of ecological analyses, we aim to aid understanding of the findings from these studies to inform future policy decisions regarding the use of ART for HIV prevention.

Introduction

Ecological studies use observational data to examine relationships between exposures and outcomes at the level of groups rather than individuals [1]. When individual-level data are unavailable, ecological studies can provide important insight into population-level trends [2,3]. Ecological studies appeal to researchers and policy-makers because they are inexpensive, use existing data, and are applicable to a broad range of issues. However, statistical models using only group-level data cannot evaluate person-level details and are therefore unable to test etiological hypotheses [2,4–6]. Further, because ecological studies often use separate data sources to measure exposures and their potential effects, the link between exposures and outcomes cannot be determined at the individual level.

Concern over these limitations has focused on "ecological fallacy," in which associations detected at the population level are mistakenly interpreted to reflect the experience of individuals in that population [1]. The first study describing ecological fallacy presented an analysis of literacy and immigration in the US, in which states with higher proportions of immigrants were shown to have higher average literacy rates [7]. An "ecologically fallacious" interpretation of this association would be that immigrants have higher literacy rates than native-born individuals; in fact, individual-level analysis shows lower literacy rates among immigrants. The best explanation for this particular population-level observation is that immigrants tend to settle in sites where the native-born individuals have higher literacy levels [7].

Despite their limitations, ecological studies play an important role in generating hypotheses that can be tested in experimental or individual-level observational studies [2,8]. For instance, ecological analyses were successfully applied during the exploratory phases of research on male circumcision to prevent HIV, in which geographical associations between circumcision rates and HIV prevalence [9–11] provided the foundation for two decades of observational research [12,13] on the topic. All three randomized clinical trials that followed were halted because of a readily demonstrable reduction of HIV acquisition in circumcised men [14,15]. A Cochrane review published in 2009 concluded that male circumcision is a clinically viable HIV prevention strategy [16].

Here we describe an illustrative set of observational studies that use ecological measures to examine the population-level effects of antiretroviral therapy (ART) on HIV transmission. We critically review what these studies measure, how they measure it, and how their findings are interpreted. These results are used to provide insight into the strengths and limitations of this approach.

Review articles synthesize in narrative form the best available evidence on a topic.

Citation: Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS (2012) HIV Treatment as Prevention: The Utility and Limitations of Ecological Observation. PLoS Med 9(7): e1001260. doi:10.1371/journal.pmed.1001260

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: MKS and KEM were supported by a National Institute of Allergy and Infectious Diseases (NIAID) T32 training grant, T32 AI0700. WCM and KAP were supported by U.S. National Institutes of Health grant R01 AI083059. MSC received funding from the University of North Carolina Center for AIDS Research, the HIV Prevention Trials Network, and R37 DK049381. The funders had no role in the study design data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: ART, antiretroviral therapy; CVL, community viral load; HPTN 052, HIV Prevention Trials Network 052; IDU, injection drug user; MSM, men who have sex with men; STARHS, serologic testing algorithm for recent HIV seroconversion.

* E-mail: mscohen@med.unc.edu

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Ecological Studies of ART for HIV Prevention

The narrative of exploring the effects of treatment on prevention shares similarities with the narrative for male circumcision, though with a somewhat different chronology. The hypothesis that antiretroviral agents can prevent sexual HIV transmission was suggested in 1988 shortly after the development of azidothymidine, which was found to effectively penetrate the genital tract [17]. This report was followed by more intensive study of the effect of newly developed antiretroviral agents on HIV replication in the male and female genital tract [18–21]. In 1994, Musicco et al. observed that azidothymidine could reduce transmission of HIV in a cohort of discordant couples by 50% [22]. Several clinical trials in the late 1990s showing that ART stopped mother-to-child transmission lent further credibility to the potential use of ART to prevent sexual transmission [23]. In 2000, a randomized clinical trial, HIV Prevention Trials Network 052 (HPTN 052), was launched to determine the magnitude and durability of the effect of combination antiretroviral agents on the prevention of sexual transmission of HIV [24]. After this trial was launched, several [25–28], but not all [29], individual-level observational studies reported a protective effect of ART against HIV transmission in serodiscordant couples. In addition, many modeling exercises suggested varying degrees of population-level prevention benefit from broader use of ART [30], the most widely discussed of which predicted elimination of HIV within five years under ideal conditions [31]. These models are discussed in a review by Eaton et al. [32] in the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections." A third group of eight ecological studies examined the population-level effects of widespread ART on HIV incidence using ecological measures, and reported significant effects [33–38] in all but two cases [39,40]. These eight studies are the focus of this review. Finally, in mid-2011, the HPTN 052 investigators reported a 96% reduction of HIV transmission in heterosexual couples over the 1.7 years of follow-up [41].

The promise of ART to control—and perhaps even eliminate [31]—HIV has mobilized calls from public health leaders to integrate preventive and clinical applications of ART [42–45]. In light of several trials showing markedly improved survival for those initiating ART earlier in the course of disease, such initiatives often emphasize the clinical benefits that early treatment can bring HIV-infected persons [46,47]. However, numerous behavioral, epidemiological, and programmatic challenges may well limit the ability to translate the individual-level prevention benefits of ART to a larger population [48–52]. As such, demonstration of a minimally biased population-level benefit is critical. Not surprisingly, there is a credible tension between the need for more randomized individual- and community-level trials (also called cluster randomized controlled trials), and the immediate scale-up of HIV treatment to prevent the spread of HIV [53–55]. The arguments for immediate and broader roll-out of ART for the sake of prevention are based on the HPTN 052 study [41], observational studies of transmission within HIV-discordant couples [25–29], ecological reports [33–36], and modeling exercises [31,56–59].

In this report we examine eight influential ecological studies that assess the population-level effects of ART on HIV transmission (Table 1). Most of the studies are from North America [33,34,36,38–40], with one set in Taiwan [35] and one in Australia [37]. Each study uses an ecological measure of the exposure, such as access to ART, or the outcome, such as HIV incidence, or both (summarized in Table 1; further considerations detailed in Table 2).

Measuring Population Exposure to ART

The simplest way investigators have characterized ART exposure in a population of HIV-infected persons is to use a dichotomous "before/after" measure, as in the case of Fang et al. [35] and Porco et al. [38], based on the time at which scale-up of local HIV treatment policies improved access to ART. Other investigators have used more detailed measures of ART exposure, including Montaner et al. [33], who estimated the number of HIV-infected persons known to be receiving ART in a population, or Katz et al. [40], who used prevalence of ART use among all identified HIV patients. How well these measures reflect actual ART exposure of an entire HIV-infected population depends on the extent to which some subpopulations remain "hidden" to investigators. ART exposure of the entire HIV-infected population can only be measured if every person with HIV infection can be identified and their treatment status assessed.

The hypothesis that population ART usage will decrease HIV incidence relies on the assumption that ongoing HIV care will sustain viral suppression, which is essential to transmission prevention [60]. However, large numbers of HIV-infected persons are lost to follow-up along the path from testing to suppressive treatment [61,62]; the US Centers for Disease Control and Prevention recently estimated that only about 24% of the 1.2 million people in the US with HIV infection in 2010 were virally suppressed (Figure 1) [61–63]. Even once in care, rates of treatment refusal by eligible individuals can be substantial [64].

To address the shortcomings of measures that do not fully reflect *suppressive* ART use, alternative metrics incorporate viral load information, based on the well-understood relationship between sexual transmission of HIV from infected individuals and viral concentrations in their blood [65] and genital fluids [66]. One such measure uses the proportion (or absolute number) of treated individuals in a study population with undetectable viral load, usually defined as having fewer than 400 copies/ml [33,37–39].

A related measure, community viral load (CVL), is used by Das et al. [34], Montaner et al. [33], and Wood et al. [36], and is defined as the total, mean [33,34], or median [36] viral load for a particular group or geographic region in a given period of time (usually a year). CVL may be a useful biomarker for describing population-level treatment outcomes over time, particularly in cases in which geospatial information about the patients' primary residence or point of medical care is available, allowing investigators to compare geographic disparities in CVL with other predictive factors such as socioeconomic status or proximity to health care programs [34,39]. However, because most CVL measures rely on public health surveillance data [33,34,39], these exposure measures reflect the treatment outcomes only for the subset of the HIV-infected population who get tested for HIV, link to care, and remain in care long enough to contribute such measurements. Patients with acute infection unidentified by serological testing are de facto not considered in the calculation of CVL, but may well be expected to contribute disproportionately to onward HIV transmission [67,68]. Additionally, the use of an aggregate measure of viral load in a community cannot capture other important drivers of HIV transmission, such as the distribution of viral loads within the population, sexual and drug-using behaviors, and the sexual or drug-use networks through which these behaviors spread HIV.

Outcome Assessment: Tracking Population HIV Transmission

Accurate assessment of HIV incidence is critical for evaluating the population-level effect of interventions, but such assessment is challenging. The simplest approach is taken by Law et al. [37], who simply refer to HIV incidence trends cited in past publications [37]. Another approach, taken by Montaner et al. [33], Das et al. [34], and Castel et al. [39], estimates population-based incidence

Table 1. Summary of exposure and outcome measures in studies using ecological measures to assess population-level effects of ART on HIV transmission.

Study (Year)	Study Location	Exposure: Trends in Population-Level Infectiousness			Outcome: Trends in HIV Transmission		
		Assessment	Measure	Trend Direction	Assessment	Measure	Trend Direction
Castel et al. [39] (2011)	Washington, D. C., US	Population in clinical care	Annual mean and total CVL; portion with undetectable VL	↓; ↑	Sentinel surveillance	Annual numbers of newly diagnosed cases	→
Das et al. [34] (2010)	San Francisco, US	Population in clinical care	Annual mean and total CVL	↓	Sentinel surveillance	Annual numbers of newly reported HIV diagnoses; annual incidence rate estimated from surveillance data using STARHS (trend was not statistically significant)	↓; ↓
Fang et al. [35] (2004)	Taiwan (national)	Time period	Time period (pre- versus post-ART period)	—	Sentinel surveillance	Surveillance data used to calculate average annual HIV transmission rate (new cases/ prevalent cases)	↓
Katz et al. [40] (2002)	San Francisco, US	Population in clinical care	Annual prevalence of ART use in HIV-infected MSM identified in HIV/AIDS registry	↑	Convenience sample of MSM	STARHS-estimated incidence using data from a VCT clinic and a STD clinic	→; ↑
Law et al. [37] (2011)	Australia (national)	Population in clinical care	Annual portion of treated patients with undetectable VL	↑	No direct assessment	Reference to previous publication describing HIV incidence trends during the same study period	↓
Montaner et al. [33] (2010)	British Columbia, Canada	Population in clinical care	Annual numbers of HIV patients receiving HAART; annual mean CVL	↑; ↓	Sentinel surveillance	New HIV-positive tests per 100 population	↓
Porco et al. [38] (2004)	San Francisco, US	Probability sample of MSM	Predicted per contact infectivity during the pre- and post-ART periods	↓	Probability sample of MSM	Annual testing in a cohort of HIV-negative MSM	↓
Wood et al. [36] (2009)	Vancouver, Canada	Convenience sample of IDUs	Biannual median CVL	↓	Convenience sample of IDUs	Annual testing in a cohort of HIV-negative IDUs	↓

↑, upward trend; ↓, downward trend; →, stable rate. For studies using two exposure or outcome measures, two arrows are shown, corresponding to the measures listed first and second.
HAART, highly active ART; STD, sexually transmitted disease; VCT, voluntary counseling and testing; VL, viral load.
doi:10.1371/journal.pmed.1001260.t001

Table 2. Summary of measures used and considerations for their use.

Measure	Considerations
Exposure: trends in population-level exposure to suppressive ART	
Before/after ART	Dichotomous measure does not quantify the level of suppressive ART use in a population
Prevalence of ART use	Only represents the prevalence of ART use among populations in clinical care; does not account for failure to achieve viral suppression
Portion of treated individuals with undetectable VL	Only represents the portion of individuals with undetectable VL among populations in clinical care
CVL	Only represents VL among populations in clinical care; aggregate and mean values obscure important differences in transmissibility among individuals
Outcome: trends in HIV transmission	
New HIV diagnoses	New HIV diagnoses do not necessarily represent incident cases.
HIV incidence via longitudinal cohort follow-up	Individuals who enroll and stay in cohorts may have lower HIV incidence than those who do not; choice of testing interval and assay can introduce bias
HIV incidence via laboratory-based methods for identifying recent infections	Recent infections identified only among HIV-infected individuals who test; assays have low specificity and can overestimate recent infections
HIV transmission rates (new cases per prevalent case-year) from modified back-calculation approach	Sensitive to peculiarities of a population's testing behavior, including frequent repeat testers or variable rates of disease progression among identified cases

VL, viral load.
doi:10.1371/journal.pmed.1001260.t002

from information on newly identified cases, using new HIV diagnoses as a direct proxy for new infections. Obviously, newly diagnosed patients acquired HIV at some unknown earlier time, and so they are not "incident" in the traditional use of the word. Using new diagnoses as a proxy for incidence also misses populations that do not seek testing and that may have lower access to health care and a corresponding higher risk for acquiring HIV [69,70]. Changes in the number of new HIV diagnoses may reflect actual changes in incidence, but will also be affected by changes in availability of services and testing behaviors [71].

A second population-based incidence estimation method, used by Fang et al. [35], back-calculates past incidence from new diagnoses [35]. This method relies in part on assumptions of uniform parameters for disease progression markers such as the onset of AIDS symptoms or the proportion of newly diagnosed

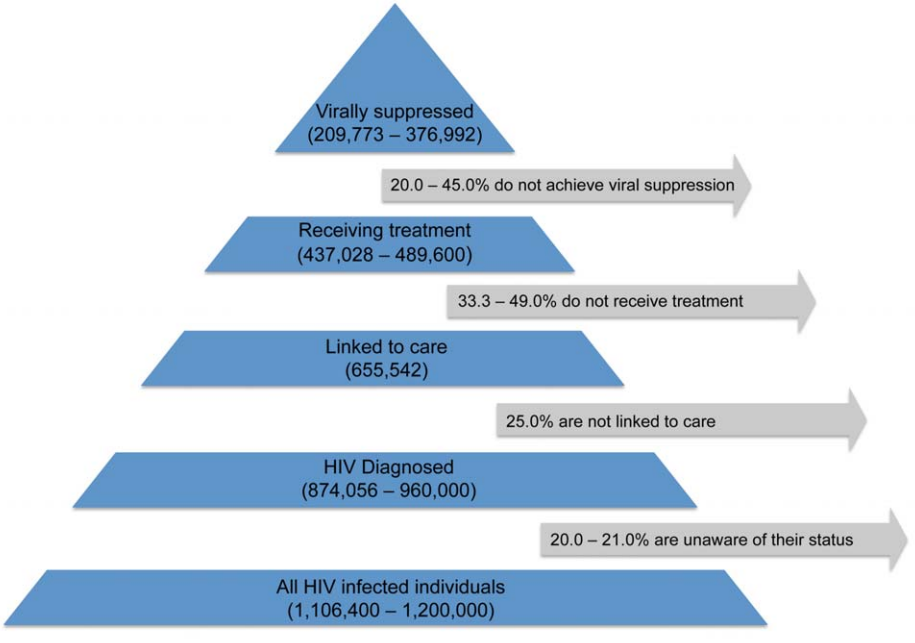

Figure 1. Estimated numbers of HIV-infected individuals in the US retained (and corresponding percentages lost) at various stages of the test, link, and treat cascade. This figure is based on data from [61,62].
doi:10.1371/journal.pmed.1001260.g001

individuals with CD4 cell count less than 200 cells/μl, although variation in the rate of decline of CD4 cell count over time and across gender, ethnicity, and HIV subtypes undermines the validity of this method [72–74].

Laboratory assays to identify persons with recent HIV infection can be applied to stored biospecimens collected in the course of routine surveillance or epidemiological research studies and may provide a more rigorous method to determine current HIV incidence from new diagnoses. The serologic testing algorithm for recent HIV seroconversion (STARHS) [75] derives current HIV incidence from the prevalence of recent infections, based on the assay window period, delineated by the seroconversion dates as detected by the original HIV-1 antibody test and the STARHS method, and adjusting for the estimated prevalence among non-testers and the probability that HIV-infected individuals will test, receive treatment, and/or have missing specimens. Although the investigations that use this method take advantage of existing surveillance data, as in the case of Das et al. [34] and Katz et al. [40], logistical challenges in storing and tracking remnant blood can affect the completeness of data. Furthermore, even relatively new laboratory methods such as the detuned enzyme-linked immunosorbent assay or the newer BED capture enzyme immunoassay have been known to misclassify established infections as incident infections; therefore, results must be interpreted with some caution [76]. The use of this approach has generally fallen out of favor pending development of better laboratory-based tests or algorithms [77].

In contrast to population-based methods, longitudinal cohort follow-up data have also been used to define population incidence, as in the analyses carried out by Wood et al. [36] and Porco et al. [38]. Although long considered the gold standard of HIV incidence estimation, cohort follow-up is not immune to bias, such as that which can result from the choice of testing intervals and HIV assay [78]. Moreover, cohort participants may be a poor proxy for the rest of the population, especially if the individuals who enroll and remain in the cohort have fewer risk behaviors than their unobserved counterparts.

Identifying the Effects of ART on HIV Transmission

The ultimate aim of these investigations is to determine whether population-level ART exposure has affected HIV transmission. Some investigators used inductive reasoning to synthesize either their own results [38,40] or a combination of their own results and other published reports [37]. The remaining studies quantify the association by comparing transmission rates, defined as the ratio of new cases to prevalent cases in an interval of time, before and after introduction of ART [35], or by using time series regression modeling [33,34,36,39]. Although these methods of analysis differ considerably, it is worth consideration that nearly every study arrives at the same conclusion: that increased population exposure to ART leads to lower HIV transmission (Table 3).

However, inaccurate assessment of exposures or outcomes can generate bias [51]. Overestimating the decline in incidence, for instance—perhaps because of an unrecognized change in testing behaviors—could produce an upward bias in the estimated impact of ART on HIV transmission. Additionally, statistical associations do not show causation, and observed trends in HIV diagnoses may be due to factors other than population-level exposure to ART. For example, declines in HIV incidence in settings worldwide—most of which started to occur before ART was available or could be expected to have had an effect—have been ascribed to various phenomena, including the saturation of HIV in high-risk groups [79] and changes in sexual behavior in response to the HIV pandemic [71,78]. Although the potential confounding effects of changes in HIV-related risk behaviors have been widely acknowledged, only one report, from Vancouver [36], formally controls for them in a regression model (Table 3). By comparison, another study from British Columbia attributes large numbers of averted HIV infections among injection drug users (IDUs) to broader uptake of ART in the community, but some have suggested that the analysis underestimates the potential protective effects of other HIV prevention measures directed at the same community [80]. Indeed, the protective effects of Vancouver's safer injection sites have been documented in the past [81,82]. Also, consistent ART adherence may be difficult to sustain in IDUs [83], further suggesting that factors beyond viral suppression may have contributed to the reduction in HIV incidence in this population.

The ability of ART to visibly reduce the number of newly diagnosed cases of HIV takes time, because most new diagnoses are made years after infection occurs, and many patients present with a reduced CD4 count, reflecting substantial progression of HIV disease. But in some ecological studies, the effect of ART is presumed to be almost immediate. In the report from British Columbia [33], where combination ART was introduced in 1996, the largest decrease in documented new HIV diagnoses took place between 1997 and 2000, but it is reasonable to question whether enough suppressive combination ART was immediately available to most patients to explain this decline.

Alternative Results and Other Considerations

The comparative lack of reports investigating the ecological effects of population-level ART in settings where rising incidence rates have been detected [84,85] suggests potential publication bias. It is also noteworthy that ecological studies of ART for HIV prevention are almost exclusively from developed western settings, likely because of the limited availability of surveillance data, viral load measurements, or registry data in resource-constrained settings.

Stable or rising HIV incidence among certain population subgroups with ready access to ART suggests the possibility that identified relationships between ART access and declines in HIV diagnoses in the studies reviewed here may be overstated. For example, HIV incidence (estimated by STARHS) increased and then stabilized among voluntary testers in San Francisco between 1999 and 2006 [85], and model-estimated numbers of new HIV infections among men who have sex with men (MSM) in British Columbia increased by 13% from 2005 to 2008 [86]. In parts of Australia, the number of HIV diagnoses among MSM between 2000 and 2006 doubled, although cohort data suggest that this observation may be largely driven by new infections among older MSM [87,88]. In Denmark [89] and the UK [90], incidence rates among MSM have reportedly increased. In Canada, some subgroups of IDUs have experienced rising HIV incidence, including Aboriginals [91], women [92], and youth [93], prompting a call for renewed prevention efforts [94].

More than 50 experimental studies of treatment as prevention are in some stage of development, and more can be anticipated [95,96]. Policy-makers often do not have the luxury of waiting years for trial data, and all decisions take place under a certain degree of uncertainty. To this end, several studies, including some considered in this review, have successfully applied novel tools of geospatial mapping and phylogenetic analysis to aid interpretation of observational data. A study in the UK [97] used viral molecular phylogeny to determine the single most likely transmitter among

Table 3. Analysis methods and conclusions regarding effects.

Author (Year)	Analysis Method	Statistical Analysis Results	Control for Confounders?	Conclusions
Castel et al. [39] (2011)	Negative binomial regression of new diagnoses on mean CVL	No effect estimate given, but lack of association reported ($p = 0.11$)	No	No association was found between trends in the mean CVL and newly diagnosed HIV/AIDS cases
Das et al. [34] (2010)	Poisson regression of new diagnoses on changes in total and mean CVL; meta-regression of estimated incidence on changes in total and mean CVL	No effect estimate given, but statistically significant trend with new diagnoses noted ($p = 0.003$)	Notes reported trends in rectal gonorrhea, but no formal assessment	Reductions in CVL were associated with a decrease in new HIV diagnoses, but not with slight HIV incidence decrease
		No effect estimate given, but lack of association with estimated incidence noted ($p > 0.30$)		
Fang et al. [35] (2004)	Modified back-calculation to estimate reduction in transmission rate (new cases per prevalent case-year) between pre- and post-HAART eras	Pre-HAART transmission rate estimated as 0.391 new infections per prevalent case	Secondary analysis of concurrent trends in annual reported cases of syphilis and gonorrhea, but no formal assessment	Provision of free ART was associated with a 53% reduction in the estimated HIV transmission rate
		Post-HAART transmission rate estimated as 0.184 new infections per prevalent case		
Katz et al. [40] (2002)	Inferences drawn from observation of concurrent changes in HIV incidence rates, reported sexual behavior, STI diagnoses, and ART use among population in clinical care	—	Secondary analysis of concurrent trends in reported risk behaviors and cases of rectal gonorrhea among MSM, but no formal assessment	ART impact on HIV transmission has been counterbalanced by increased reported risk behaviors
Law et al. [37] (2011)	Inferences drawn from predicted changes in prevalence of undetectable VL among population in clinical care and external reports of HIV incidence	—	No	Declines in predicted detectable VL between 1997 and 2009 coincide with reports of rising new diagnoses and estimated incidence in the same community
Montaner et al. [33] (2010)	Poisson regression of estimated new diagnoses on changes in median CVL and numbers receiving HAART	Effect of 100 new patients receiving HAART on estimated new diagnoses predicted as −0.97 (95% CI 0.96–0.98)	Notes reported trends in infectious syphilis, rectal gonorrhea, and genital chlamydia as proxies for sexual risk behaviors; trends in hepatitis C were also noted as proxy for unsafe injecting behaviors	Increased ART coverage and reduced CVL are associated with a decreased number of new HIV diagnoses
		Effect of 1 log decrease median CVL on estimated new diagnoses predicted as −0.86 (0.75–0.98)		
Porco et al. [38] (2004)	Inferences drawn from trends in annual HIV incidence based on antibody testing and time period (pre- versus post-HAART period) as indicator of ART use	—	Transmission probability accounts for sexual risk behaviors among surveyed MSM	Wider availability of ART appears to have slowed transmission in the study population
Wood et al. [36] (2009)	Unadjusted and adjusted Cox proportional hazards regression of time to seroconversion on median CVL in the preceding six months	Unadjusted hazard ratio for effect of median CVL on time to seroconversion estimated as 3.57 (2.03–6.27) per $\log_{10}$ CVL increase	Adjusted model controlled for needle sharing, unprotected sex, ethnicity, daily heroin use, and unstable housing	Median CVL predicts HIV incidence independent of HIV risk behaviors
		Adjusted hazard ratio for effect of median CVL on time to seroconversion estimated as 3.32 (1.82–6.08) per $\log_{10}$ CVL increase		

CI, confidence interval; HAART, highly active ART; STI, sexually transmitted infection; VL, viral load.
doi:10.1371/journal.pmed.1001260.t003

MSM, allowing the investigators to account for the higher transmission probability of individuals with acute and early infection. Other studies, including those from San Francisco [34] and Washington, D. C. [39], used geospatial analysis to illustrate the spatial distribution of HIV-infected individuals in communities. Most recently, investigators at the Africa Centre for Health and Population Studies in South Africa have been able to identify a relationship between the density of ART use and HIV acquisition risk within a community by studying HIV incidence in a longitudinal cohort of more than 16,000 individuals (personal communication, F. Tanser). Additional strengths of this study include the use of information about the patients' primary residence and attempts to control for at least some possible confounders of the relationship between ART uptake and HIV incidence in the same community.

Several large cluster randomized controlled trials are being developed [98]. A team from the Harvard School of Public Health AIDS Initiative, working with partners in Botswana, will target

individuals with sustained, high plasma viral loads for immediate treatment, a strategy that could have exponential public health benefits [73]. A second group from the London School of Hygiene & Tropical Medicine and Imperial College London plans to test the feasibility and impact of a universal test-and-treat strategy along with other combination prevention measures, including male circumcision [99]. But clinical trials have their own limitations, including time, cost, ethical challenges, and perturbations to the underlying community that can cause bias. And there is never a guarantee that approaches employed in a trial will prove effective outside of the trial setting.

Conclusions

Suppressive ART prevents HIV transmission in stable, monogamous, heterosexual couples. While ART seems to hold great promise as a public health tool, its population-level benefits have not been proven. Although ecological studies can play an important role in the development of new HIV prevention strategies, they are methodologically limited to building justification of further formal scientific inquiry into population-level effects of the potential policies in question. They are therefore the first of many steps in the path from science to policy, beginning with the establishment of biological plausibility, and progressing to assessment of an individual-level effect and then a group-level effect. Though most policy decisions must be made under conditions of uncertainty, the hypothesis that widespread ART can eliminate HIV infection [31,100] may have raised expectations beyond what can actually be achieved. Additionally, implementation of treatment as prevention is not without its risks, including the rise of population-level drug resistance with the rapid uptake of ART in the face of continued limited infrastructure, and increased risk compensation by treated individuals who believe that treatment alone may justify forgoing other forms of protection [101–103].

Although we expect an impact of ART at the population level, the magnitude of the effect may not be as great as some hope; measuring the impact of ART roll-out on HIV spread, as in several planned cluster randomized controlled trials, therefore remains a critical step. Much as combination prevention methods are believed to be better than single interventions for HIV prevention [104], all the methods available to determine the benefits of prevention interventions, including ecological studies, should be deployed. The results must be weighed and used with a full understanding of the methods used to define the outcomes of treatment of HIV infection for prevention of transmission.

Key Points

- Several strong observational studies and one randomized controlled trial, HPTN 052, demonstrate that ART reduces the transmission of HIV in stable, heterosexual HIV-discordant couples.
- A number of ecological studies, which use observational data to examine relationships between exposures and outcomes at the level of groups instead of individuals, have found associations between the broader use of antiretroviral agents and reductions in new HIV diagnoses in at-risk populations or in the general population. Ecological studies may generate hypotheses that can be explored using other experimental or observational methods.
- A better understanding of the strengths and limitations of the various exposure and outcome measures used in ecological studies that examine the population-level effects of ART on HIV transmission, and the methods used to analyze them, is essential for the effective application of the findings in policy-making processes.
- Methodological challenges such as measurement error, selection bias, confounding, and assumptions about the time lag of effects must be taken into account when interpreting the results of these studies.
- Prospective measurement of the population-level impact of ART can be approached through cluster randomized controlled trials, but the cost, time, and degree of difficulty of designing and conducting these studies are appreciable, and the potential for the intervention itself to introduce bias can threaten the validity of the results.
- Measuring the population-level benefits of ART is critical to HIV prevention efforts, and consideration of results of all methods may be used to inform ongoing research and public health policy. A firm understanding of the strengths and weaknesses of each approach is crucial to the interpretation of results and allocation of resources.

Author Contributions

Conceived and designed the experiments: MKS KAP KEM WCM MSC. Wrote the first draft of the manuscript: MKS. Contributed to the writing of the manuscript: MKS KAP KEM WCM MSC. ICMJE criteria for authorship read and met: MKS KAP KEM WCM MSC. Agree with manuscript results and conclusions: MKS KAP KEM WCM MSC.

References

1. Rothman KJ, Greenland S, Lash TL (2008) Modern epidemiology. Philadelphia: Lippincott Williams & Wilkins.
2. Piantadosi S, Byar DP, Green SB (1988) The ecological fallacy. Am J Epidemiol 127: 893–904.
3. Hallett TB, White PJ, Garnett GP (2007) Appropriate evaluation of HIV prevention interventions: from experiment to full-scale implementation. Sex Transm Infect 83 (Suppl 1): i55–i60. doi:10.1136/sti.2006.023663
4. Morgenstern H (1995) Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health 16: 61–81. doi:10.1146/annurev.pu.16.050195.000425.bv
5. Greenland S, Robins J (1994) Invited commentary: ecologic studies—biases, misconceptions, and counterexamples. Am J Epidemiol 139: 747–760.
6. Schwartz S (1994) The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health 84: 819–824. doi:10.2105/AJPH.84.5.819
7. Robinson W (1950) Ecological correlations and the behavior of individuals. Am Sociol Rev 15: 351–357. doi:10.1093/ije/dyn357
8. Durrheim D, Ogunbanjo G (2004) Making sense of statistics for the family practitioner: "what are ecological studies?" S Afr Fam Pract 46: 48.
9. Moses S, Bradley JE, Nagelkerke NJ, Ronald AR, Ndinya-Achola JO, et al. (1990) Geographical patterns of male circumcision practices in Africa: association with HIV seroprevalence. Int J Epidemiol 19: 693–697.
10. Bongaarts J, Reining P, Way P, Conant F (1989) The relationship between male circumcision and HIV infection in African populations. AIDS 3: 373–377.
11. Caldwell JC, Caldwell P (1996) The African AIDS epidemic. Sci Am 274: 62–63, 66–68.
12. Weiss HA, Quigley MA, Hayes RJ (2000) Male circumcision and risk of HIV infection in sub-Saharan Africa: a systematic review and meta-analysis. AIDS 14: 2361–2370.
13. Siegfried N, Muller M, Deeks J, Volmink J, Egger M, et al. (2005) HIV and male circumcision—a systematic review with assessment of the quality of studies. Lancet Infect Dis 5: 165–173. doi:10.1016/S1473-3099(05)01309-5
14. Bailey RC, Plummer FA, Moses S (2001) Male circumcision and HIV prevention: current knowledge and future research directions. Lancet Infect Dis 1: 223–231. doi:10.1016/S1473-3099(01)00117-7
15. Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, et al. (2007) Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet 369: 657–666. doi:10.1016/S0140-6736(07)60313-4

16. Siegfried N, Muller M, Deeks JJ, Volmink J (2009) Male circumcision for prevention of heterosexual acquisition of HIV in men. Cochrane Database Syst Rev 2009: CD003362. doi:10.1002/14651858.CD003362.pub2
17. Henry K, Chinnock BJ, Quinn RP, Fletcher CV, de Miranda P, et al. (1988) Concurrent zidovudine levels in semen and serum determined by radioimmunoassay in patients with AIDS or AIDS-related complex. JAMA 259: 3023–3026.
18. Dumond JB, Yeh RF, Patterson KB, Corbett AH, Jung BH, et al. (2007) Antiretroviral drug exposure in the female genital tract: implications for oral pre- and post-exposure prophylaxis. AIDS 21: 1899–1907. doi:10.1097/QAD.0b013e328270385a
19. Nicol MR, Kashuba ADM (2010) Pharmacologic opportunities for HIV prevention. Clin Pharmacol Ther 88: 598–609. doi:10.1038/clpt.2010.189
20. Pereira AS, Kashuba AD, Fiscus SA, Hall JE, Tidwell RR, et al. (1999) Nucleoside analogues achieve high concentrations in seminal plasma: relationship between drug concentration and virus burden. J Infect Dis 180: 2039–2043. doi:10.1086/315149
21. Kashuba AD, Dyer JR, Kramer LM, Raasch RH, Eron JJ, et al. (1999) Antiretroviral-drug concentrations in semen: implications for sexual transmission of human immunodeficiency virus type 1. Antimicrob Agents Chemother 43: 1817–1826.
22. Musicco M, Lazzarin A, Nicolosi A, Gasparini M, Costigliola P, et al. (1994) Antiretroviral treatment of men infected with human immunodeficiency virus type 1 reduces the incidence of heterosexual transmission. Italian Study Group on HIV Heterosexual Transmission. Arch Intern Med 154: 1971–1976.
23. Siegfried N, van der Merwe L, Brocklehurst P, Sint TT (2011) Antiretrovirals for reducing the risk of mother-to-child transmission of HIV infection. Cochrane Database Syst Rev 2011: CD003510. doi:10.1002/14651858.CD003510.pub3
24. Cohen MS, McCauley M, Gamble TR (2012) HIV treatment as prevention and HPTN 052. Curr Opin HIV AIDS 7: 99–105. doi:10.1097/COH.0b013e32834f5cf2
25. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, et al. (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet 375: 2092–2098. doi:10.1016/S0140-6736(10)60705-2
26. Bunnell R, Ekwaru JP, Solberg P, Wamai N, Bikaako-Kajura W, et al. (2006) Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda. AIDS 20: 85–92.
27. Sullivan P, Kayitenkore K, Chomba E, Karita E, Mwananyanda L, et al. (2010) Reduction of HIV transmission risk while prescribed anti-retroviral therapy (ARVT): Misclassification of ARVT status as a methodological issue [abstract]. AIDS Vaccine 2010; 28 Sept–1 Oct 2010; Atlanta, Georgia, US. AIDS Res Hum Retroviruses 26: A-1–A-184. doi:10.1089/aid.2010.9998
28. Del Romero J, Castilla J, Hernando V, Rodriguez C, Garcia S (2010) Combined antiretroviral treatment and heterosexual transmission of HIV-1: cross sectional and prospective cohort study. BMJ 340: c2205–c2205. doi:10.1136/bmj.c2205
29. Wang L, Ge Z, Luo J, Shan D, Gao X, et al. (2010) HIV transmission risk among serodiscordant couples: a retrospective study of former plasma donors in Henan, China. J Acquir Immune Defic Syndr 55: 232–238. doi:10.1097/QAI.0b013e3181e9b6b7
30. Gay CL, Cohen MS (2008) Antiretrovirals to prevent HIV infection: pre- and postexposure prophylaxis. Curr Infect Dis Rep 10: 323–331.
31. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57. doi:10.1016/S0140-6736(08)61697-9
32. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
33. Montaner JS, Lima VD, Barrios R, Yip B, Wood E, et al. (2010) Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 376: 532–539. doi:10.1016/S0140-6736(10)60936-1
34. Das M, Chu PL, Santos G-M, Scheer S, Vittinghoff E, et al. (2010) Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. PLoS ONE 5: e11068. doi:10.1371/journal.pone.0011068
35. Fang C-T, Hsu H-M, Twu S-J, Chen M-Y, Chang Y-Y, et al. (2004) Decreased HIV transmission after a policy of providing free access to highly active antiretroviral therapy in Taiwan. J Infect Dis 190: 879–885. doi:10.1086/422601
36. Wood E, Kerr T, Marshall BDL, Li K, Zhang R, et al. (2009) Longitudinal community plasma HIV-1 RNA concentrations and incidence of HIV-1 among injecting drug users: prospective cohort study. BMJ 338: b1649. doi:10.1136/bmj.b1649
37. Law MG, Woolley I, Templeton DJ, Roth N, Chuah J, et al. (2011) Trends in detectable viral load by calendar year in the Australian HIV observational database. J Int AIDS Soc 14: 10. doi:10.1186/1758-2652-14-10
38. Porco TC, Martin JN, Page-Shafer KA, Cheng A, Charlebois E, et al. (2004) Decline in HIV infectivity following the introduction of highly active antiretroviral therapy. AIDS 18: 81–88.
39. Castel AD, Befus M, Willis S, Griffin A, West T, et al. (2011) Use of the community viral load as a population-based biomarker of HIV burden, Washington, DC. AIDS 26: 345–353. doi:10.1097/QAD.0b013e32834de5fe
40. Katz MH, Schwarcz SK, Kellogg TA, Klausner JD, Dilley JW, et al. (2002) Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco. Am J Public Health 92: 388–394.
41. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505. doi:10.1056/NEJMoa1105243
42. 6th IAS Conference on HIV Pathogenesis, Treatment and Prevention Conference Programme (2011) Fauci: looking to Rome, with optimism. Available: http://blog.ias2011.org/post/2011/07/10/Looking-to-Rome-with-Optimism.aspx. Accessed 26 December 2011.
43. del Rio C (2011) Treatment is prevention: the results of HPTN 052. Journal Watch. Available: http://aids-clinical-care.jwatch.org/cgi/content/full/2011/719/1. Accessed 25 November 2011.
44. Squires KE, Hughes JM (2011) [Letter to the President of the United States on behalf of the Infectious Diseases Society of America and the HIV Medicine Association.] Available: http://www.hivma.org/uploadedFiles/HIVMA/Policy_and_Advocacy/Policy_Priorities/Evidence_Based_Prevention/Comments/HPTN_052_Letter.pdf#search = %22squires hughes%22. Accessed 3 June 2012.
45. Maugh TH 2nd (2011 May 13) Anti-HIV drugs prove highly effective in preventing transmission of virus. Los Angeles Times. Available: http://articles.latimes.com/2011/may/13/health/la-he-aids-prevention-20110513. Accessed 29 May 2012.
46. Panel on Antiretroviral Guidelines for Adults and Adolescents (2012) Guidelines for the use of antiretroviral agents in HIV-1 infected adults and adolescents. US Department of Health and Human Services. Available: http://aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf. Accessed 29 May 2012.
47. World Health Organization (2011) The strategic use of antiretrovirals for treatment and prevention of HIV. Geneva: World Health Organization. Available: http://www.who.int/hiv/pub/meetingreports/consultation_20111116/en/index.html. Accessed 29 May 2012.
48. Epstein H, Morris M (2011) HPTN 052 and the future of HIV treatment and prevention. Lancet 378: 225. doi:10.1016/S0140-6736(11)61117-3
49. Cates W (2011) HPTN 052 and the future of HIV treatment and prevention. Lancet 378: 224–225. doi:10.1016/S0140-6736(11)61116-1
50. Anglemyer A, Rutherford GW, Egger M, Siegfried N (2011) Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. Cochrane Database Syst Rev 2011: CD009153. doi:10.1002/14651858.CD009153
51. Schwartländer B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041. doi:10.1016/S0140-6736(11)60702-2
52. Geffen N (2011) When to start antiretroviral therapy in adults: the results of HPTN 052 move us closer to a "test-and-treat" policy. South Afr J HIV Med 12: 9–11. Available: http://sajhivmed.org.za/index.php/sajhivmed/article/view/733/586. Accessed 29 May 2012.
53. Montaner JSG (2011) Treatment as prevention—a double hat-trick. Lancet 378: 208–209. doi:10.1016/S0140-6736(11)60821-0
54. (2011) HIV treatment as prevention—it works. Lancet 377: 1719. doi:10.1016/S0140-6736(11)60713-7
55. Seale A, Lazarus JV, Grubb I, Fakoya A, Atun R (2011) HPTN 052 and the future of HIV treatment and prevention. Lancet 378: 226–225. doi:10.1016/S0140-6736(11)61118-5
56. Law MG, Prestage G, Grulich A, Van de Ven P, Kippax S (2001) Modelling the effect of combination antiretroviral treatments on HIV incidence. AIDS 15: 1287–1294.
57. Blower S, Bodine E, Kahn J, McFarland W (2005) The antiretroviral rollout and drug-resistant HIV in Africa: insights from empirical data and theoretical models. AIDS 19: 1–14.
58. Lima VD, Johnston K, Hogg RS, Levy AR, Harrigan PR, et al. (2008) Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. J Infect Dis 198: 59–67. doi:10.1086/588673
59. Meyer-Rath G, Over M (2012) HIV treatment as prevention: modelling the cost of antiretroviral treatment—state of the art and future directions. PLoS Med 9: e1001247. doi:10.1371/journal.pmed.1001247
60. Cohen MS, Gay CL (2010) Treatment to prevent transmission of HIV-1. Clin Infect Dis 50: S85–S95. doi:10.1086/651478
61. Gardner EM, McLees MP, Steiner JF, del Rio C, Burman WJ (2011) The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis 52: 793–800. doi:10.1093/cid/ciq243
62. Burns DN, Dieffenbach CW, Vermund SH (2010) Rethinking prevention of HIV type 1 infection. Clin Infect Dis 51: 725–731. doi:10.1086/655889
63. US Centers for Disease Control and Prevention (2011 Nov 29) Only one quarter of Americans with HIV have virus under control. Available: http://www.cdc.gov/nchhstp/newsroom/WAD2011PressRelease.html. Accessed 29 May 2012.

64. Katz IT, Essien T, Marinda ET, Gray GE, Bangsberg DR, et al. (2011) Antiretroviral therapy refusal among newly diagnosed HIV-infected adults. AIDS 25: 2177–2181. doi:10.1097/QAD.0b013e32834b6464
65. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med 342: 921–929. doi:10.1056/NEJM200003303421303
66. Baeten JM, Kahle E, Lingappa JR, Coombs RW, Delany-Moretlwe S, et al. (2011) Genital HIV-1 RNA predicts risk of heterosexual HIV-1 transmission. Sci Transl Med 3: 77ra29. doi:10.1126/scitranslmed.3001888
67. Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, et al. (2011) The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet 378: 256–268. doi:10.1016/S0140-6736(11)60842-8
68. Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, et al. (2007) High rates of forward transmission events after acute/early HIV-1 infection. J Infect Dis 195: 951–959. doi:10.1086/512088
69. Lopez-Quintero C, Shtarkshall R, Neumark YD (2005) Barriers to HIV-testing among Hispanics in the United States: analysis of the National Health Interview Survey, 2000. AIDS Patient Care STDS 19: 672–683. doi:10.1089/apc.2005.19.672
70. Spielberg F, Branson BM, Goldbaum GM, Lockhart D, Kurth A, et al. (2003) Overcoming barriers to HIV testing: preferences for new strategies among clients of a needle exchange, a sexually transmitted disease clinic, and sex venues for men who have sex with men. J Acquir Immune Defic Syndr 32: 318–327.
71. Shelton JD, Cohen M, Barnhart M, Hallett T (2010) Is antiretroviral therapy modifying the HIV epidemic? Lancet 376: 1824–1825. doi:10.1016/S0140-6736(10)62163-0
72. Wand H, Wilson D, Yan P, Gonnermann A, McDonald A, et al. (2009) Characterizing trends in HIV infection among men who have sex with men in Australia by birth cohorts: results from a modified back-projection method. J Int AIDS Soc 12: 19. doi:10.1186/1758-2652-12-19
73. Novitsky V, Wang R, Bussmann H, Lockman S, Baum M, et al. (2010) HIV-1 subtype C-infected individuals maintaining high viral load as potential targets for the ''test-and-treat'' approach to reduce HIV transmission. PLoS ONE 5: e10148. doi:10.1371/journal.pone.0010148
74. Wolbers M, Babiker A, Sabin C, Young J, Dorrucci M, et al. (2010) Pretreatment CD4 cell slope and progression to AIDS or death in HIV-infected patients initiating antiretroviral therapy—the CASCADE collaboration: a collaboration of 23 cohort studies. PLoS Med 7: e1000239. doi:10.1371/journal.pmed.1000239
75. Janssen RS, Satten GA, Stramer SL, Rawal BD, O'Brien TR, et al. (1998) New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. JAMA 280: 42–48.
76. Le Vu S, Pillonel J, Semaille C, Bernillon P, Le Strat Y, et al. (2008) Principles and uses of HIV incidence estimation from recent infection testing—a review. Euro Surveill 13: 18969.
77. Busch MP, Pilcher CD, Mastro TD, Kaldor J, Vercauteren G, et al. (2010) Beyond detuning: 10 years of progress and new challenges in the development and application of assays for HIV incidence estimation. AIDS 24: 2763–2771. doi:10.1097/QAD.0b013e32833f1142
78. Joint United Nations Programme on HIV/AIDS (2010) Methods for estimating HIV incidence. Epi Alert, UNAIDS Quarterly Update on HIV Epidemiology. Available: http://www.unaids.org/en/media/unaids/contentassets/documents/dataanalysis/epi_alert_2010Q1_en.pdf. Accessed 3 June 2012.
79. Shelton JD, Halperin DT, Wilson D (2006) Has global HIV incidence peaked? Lancet 367: 1120–1122. doi:10.1016/S0140-6736(06)68436-5
80. Grulich AE, Wilson DP (2010) Is antiretroviral therapy modifying the HIV epidemic? Lancet 376: 1824. doi:10.1016/S0140-6736(10)62162-9
81. Kerr T, Small W, Buchner C, Zhang R, Li K, et al. (2010) Syringe sharing and HIV incidence among injection drug users and increased access to sterile syringes. Am J Public Health 100: 1449–1453. doi:10.2105/AJPH.2009.178467
82. Stoltz J-A, Wood E, Small W, Li K, Tyndall M, et al. (2007) Changes in injecting practices associated with the use of a medically supervised safer injection facility. J Public Health (Oxf) 29: 35–39. doi:10.1093/pubmed/fdl090
83. Wood E, Montaner JSG, Yip B, Tyndall MW, Schechter MT, et al. (2003) Adherence and plasma HIV RNA responses to highly active antiretroviral therapy among HIV-1 infected injection drug users. CMAJ 169: 656–661.
84. Murray JM, Prestage G, Grierson J, Middleton M, McDonald A (2011) Increasing HIV diagnoses in Australia among men who have sex with men correlated with the growing number not taking antiretroviral therapy. Sex Health 8: 304–310. doi:10.1071/SH10114
85. Truong HM, Kellogg T, Klausner JD, Katz MH, Dilley J, et al. (2006) Increases in sexually transmitted infections and sexual risk behavior without a concurrent increase in HIV incidence among men who have sex with men in San Francisco: a suggestion of HIV serosorting? Sex Transm Infect 82: 461–466. doi:10.1136/sti.2006.019950
86. British Columbia Centre for Disease Control (2011) HIV and sexually transmitted infections 2010. Available: http://www.bccdc.ca/NR/rdonlyres/2035512C-DBEC-495B-A332-C410EE9520C7/0/CPS_Report_STI_HIV_2010_annual_report_FINAL_20111122.pdf. Accessed 29 May 2012.
87. Guy RJ, McDonald AM, Bartlett MJ, Murray JC, Giele CM, et al. (2007) HIV diagnoses in Australia: diverging epidemics within a low-prevalence country. Med J Aust 187: 437–440.
88. Grulich AE, Kaldor JM (2008) Trends in HIV incidence in homosexual men in developed countries. Sex Health 5: 113–118. doi:10.1071/SH07075.
89. Dukers NH, Spaargaren J, Geskus RB, Beijnen J, Coutinho RA, et al. (2002) HIV incidence on the increase among homosexual men attending an Amsterdam sexually transmitted disease clinic: using a novel approach for detecting recent infections. AIDS 16: F19–F24.
90. Dougan S, Elford J, Chadborn TR, Brown AE, Roy K, et al. (2006) Does the recent increase in HIV diagnoses among men who have sex with men in the UK reflect a rise in HIV incidence or increased uptake of HIV testing? Sex Transm Infect 83: 120–125. doi:10.1136/sti.2006.021428
91. Duncan KC, Reading C, Borwein AM, Murray MC, Palmer A, et al. (2011) HIV incidence and prevalence among aboriginal peoples in Canada. AIDS Behav 15: 214–227. doi:10.1007/s10461-010-9792-y
92. Spittal PM, Craib KJP, Wood E, Laliberté N, Li K, et al. (2002) Risk factors for elevated HIV incidence rates among female injection drug users in Vancouver. CMAJ 166: 894–899.
93. Miller CL, Wood E, Spittal PM, Li K, Frankish JC, et al. (2004) The future face of coinfection: prevalence and incidence of HIV and hepatitis C virus coinfection among young injection drug users. J Acquir Immune Defic Syndr 36: 743–749.
94. Bruneau J, Daniel M, Abrahamowicz M, Zang G, Lamothe F, et al. (2011) Trends in human immunodeficiency virus incidence and risk behavior among injection drug users in montreal, Canada: a 16-year longitudinal study. Am J Epidemiol 173: 1049–1058. doi:10.1093/aje/kwq479
95. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469.
96. Boily MC, Buvé A, Baggaley RF (2010) HIV transmission in serodiscordant heterosexual couples. BMJ 340: c2449. doi:10.1136/bmj.c2449
97. Fisher M, Pao D, Brown AE, Sudarshi D, Gill ON, et al. (2010) Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. AIDS 24: 1739–1747. doi:10.1097/QAD.0b013e32833ac9e6
98. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250
99. Imperial College London Faculty of Medicine (2011) About PopART. Available: http://www1.imperial.ac.uk/medicine/research/researchthemes/infection/infectious_diseases/hiv_trials/hiv_prevention_technologies/popart/about/. Accessed 3 June 2012.
100. (2011 Jun 2) The end of AIDS? The Economist. Available: http://www.economist.com/node/18774722. Accessed 29 May 2012.
101. Crepaz N, Lyles CM, Wolitski RJ, Passin WF, Rama SM, et al. (2006) Do prevention interventions reduce HIV risk behaviors among people living with HIV? A meta-analytic review of controlled trials. AIDS 20: 143–157. doi:10.1097/01.aids.0000196166.48518.a0
102. Hasse B, Ledergerber B, Hirschel B, Vernazza P, Glass TR, et al. (2010) Frequency and determinants of unprotected sex among HIV-infected persons: the Swiss HIV Cohort Study. Clin Infect Dis 51: 1314–1322. doi:10.1086/656809
103. Cohen MS (2010) HIV treatment as prevention and ''the Swiss statement'': in for a dime, in for a dollar? Clin Infect Dis 51: 1323–1324. doi:10.1086/656810
104. Coates TJ, Richter L, Caceres C (2008) Behavioral strategies to reduce HIV transmission: how to make them work better. Lancet 372: 669–684. doi:10.1016/S0140-6736(08)60886-7

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Natural Experiments Highlight Limits of Antiretroviral Treatment as HIV Prevention

David P. Wilson*

The Kirby Institute, Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia

Abstract: There is growing enthusiasm for increasing coverage of antiretroviral treatment among HIV-infected people for the purposes of preventing ongoing transmission. Treatment as prevention will face a number of barriers when implemented in real world populations, which will likely lead to the effectiveness of this strategy being lower than proposed by optimistic modelling scenarios or ideal clinical trial settings. Some settings, as part of their prevention and treatment strategies, have already attained rates of HIV testing and use of antiretroviral therapy—with high levels of viral suppression—that many countries would aspire to as targets for a treatment-as-prevention strategy. This review examines a number of these "natural experiments", namely, British Columbia, San Francisco, France, and Australia, to provide commentary on whether treatment as prevention has worked in real world populations. This review suggests that the population-level impact of this strategy is likely to be considerably less than as inferred from ideal conditions.

Introduction

HIV prevention decision-makers across the world are considering the expansion of antiretroviral therapy (ART) for HIV-infected people in order to reduce their infectiousness and thus prevent onward transmission. This approach, called treatment as prevention, is a paradigm shift from using ART for the sole purpose of improving the health and longevity of patients with HIV. We are now in an era where the secondary benefit of ART is being considered as potentially the primary public health approach to controlling HIV epidemics.

Several findings suggest that treatment might be effective as prevention: the HPTN 052 study demonstrated that ART reduces sexual transmission between discordant couples in a trial setting [1]; various ecological studies from community settings have shown an association between ART programs and reduced markers of incidence [2–5]; associations have been demonstrated between reduced viral load and lower infectiousness [6–8]; and some theoretical models even suggest that under idealised conditions, elimination might be possible [9,10]. However, these findings do not imply that widespread scale-up of ART programs under real world conditions will reduce HIV incidence at a population level to the degree that some people are expecting (i.e., towards elimination). Cluster-randomised trials are currently underway in Africa to investigate the impact of high coverage of ART at the population level. In the meantime, models are projecting potential epidemic trajectories associated with treatment-as-prevention strategies under less ideal conditions [11], and various national and international organisations are already discussing operational issues about how to implement treatment as prevention [12].

We do not need to wait for trials of increased ART coverage to be completed, or speculate through the use of mathematical models, to have some understanding of the likely population-level impact of this strategy. Treatment as prevention has essentially been implemented in some settings already for a considerable time. Planned treatment-as-prevention approaches involve frequent universal testing and initiation of ART early post-diagnosis, but increasing treatment coverage at any stage of infection—and reaching high degrees of viral suppression across a population of people living with HIV—is de facto treatment as prevention. Some settings have achieved these objectives as part of their independent prevention and treatment responses: these settings can be considered as natural experiments for treatment as prevention at the population level.

Natural Experiment Case Studies

British Columbia, Canada

A study by Montaner et al. [3] has been widely promoted as demonstrating treatment as prevention in a community setting, namely, among people who inject drugs (PWID) in British Columbia, Canada. In British Columbia, there is universal access to free rapid HIV testing (though it is not known what proportion of PWID get tested for HIV each year). Guidelines for ART in British Columbia indicate that any HIV-positive patient may commence treatment, regardless of CD4 count, and ART is recommended for all symptomatic patients with established disease, and for asymptomatic individuals with CD4 cell count $\leq$500 cells/μl [13]. Estimates for ART coverage are difficult to quantify precisely, but coverage is considered to be relatively high and has certainly increased over time.

Citation: Wilson DP (2012) HIV Treatment as Prevention: Natural Experiments Highlight Limits of Antiretroviral Treatment as HIV Prevention. PLoS Med 9(7): e1001231. doi:10.1371/journal.pmed.1001231

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: No specific funding was received for writing this article.

Competing Interests: The author has declared that no competing interests exist.

Abbreviations: ART, antiretroviral therapy; MSM, men who have sex with men; PWID, people who inject drugs

* E-mail: dwilson@unsw.edu.au

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Montaner et al. found that there was an association between declining rates of HIV diagnoses and increasing rates of testing, ART coverage, and viral suppression. However, it is not clear to what extent the reduction in incidence is attributable to ART versus other interventions. As discussed by Smith et al. [14], also in the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections", analyses conducted for British Columbia have been ecological, and declines in incidence could be attributed to other prevention programs specifically targeting this population group over the same period [15].

San Francisco, United States

With high, and increasing, rates of HIV testing and ART coverage and effectiveness, San Francisco is an obvious case study for evaluating the role of treatment as prevention. It is estimated that rates of HIV testing have been increasing in San Francisco, such that ~72% of the core group at risk of infection, namely, men who have sex with men (MSM), received an HIV test in the past 12 months, and only 15%–20% of HIV cases are undiagnosed [4]. An increasing proportion of HIV-infected patients are enrolled in care (~71% in 2004 and 78% in 2008) [16], and mean levels of community viral load have significantly decreased (from ~25,000 copies/ml in 2004/2005 to 15,000 copies/ml in 2008) [4].

Das et al. [4] used San Francisco's HIV/AIDS surveillance system to examine trends in community viral load and new HIV diagnoses, as a surrogate marker for incidence. They found that reductions in community viral load were associated with decreased HIV diagnoses since 2004. As a purely ecological study, causation cannot be attributed to ART, but their results suggest that high coverage of ART could have reduced HIV transmission at the population level. However, although the number of newly diagnosed and reported HIV cases has been declining in San Francisco, the rate of new infections is still relatively high [4], possibly because of the substantial HIV prevalence (~25%) among MSM [17,18]. As such, even if the average per individual infectiousness is reduced, there is still likely to be a significant number of new HIV infections occurring at the population level each year.

France

A "treatment as prevention" statement has been released by the French National AIDS Council [19], which takes a less assertive approach to "test and treat" but still strongly promotes testing and treatment. The level of undiagnosed infections in France is approximately 25%–30%, comparable to levels in other resource-rich settings [20–28]. ART guidelines in France in 2007 indicated that ART should be started as early as possible for symptomatic patients and those with high viral loads (>100,000 copies/ml), and for asymptomatic patients when the CD4 count reached 350 cells/μl [29]. There have been significant increases in the uptake of ART among eligible people in France (to ~85%), and ~92% of treated patients achieve plasma viral suppression [30]. However, treatment-as-prevention strategies cannot be said to have been fully implemented in France, as many patients initiate ART too late [29].

The outcomes of the natural experiment in France suggest that there may be differences between at-risk groups in the population-level effectiveness of ART for reducing incidence: HIV incidence has declined or remained stable in all major population groups, except MSM, where incidence has been high and increasing [30]. Data from behavioural studies indicate that unprotected anal sex and numbers of sexual partners among MSM have increased [31] (also coinciding with increases in syphilis transmission [32]), raising the possibility that disinhibition or independent sociobehavioural changes could undercut the effectiveness of treatment as prevention. It is also possible that the increased HIV incidence among MSM could be due to higher risk behaviours among those who are not on ART and do not have suppressed virus.

Australia

Australia could also be considered a setting where a natural experiment for treatment as prevention has taken place. First, the HIV epidemic is highly concentrated, with the majority (~80%) of all HIV cases being among MSM [33], a population generally well educated and actively engaged with respect to HIV. Second, HIV testing is routinely carried out by most MSM, with approximately 60%–75% of men self-reporting an annual test [34] and just 10% of men reporting having never been tested [35]. Third, all regimen lines and combinations of ART are publicly funded and freely available to all HIV-infected patients. Australian guidelines for treatment advise considering ART when CD4 cell count is <500 cells/μl, and definitely treating when CD4 cell count is <350 cells/μl. There are increased numbers of people receiving ART in Australia (at about 70% of all individuals living with HIV) [36]; however, about 20% of individuals commence ART when CD4 cell count is <200 cells/μl, because of late presentation [33]. Fourth, the proportion of people on ART with undetectable viral load has increased from 65% to 90% (at 400 copies/ml sensitivity; 40% to 85% at 50 copies/ml sensitivity) [37]. Further information and analyses on these data are provided elsewhere [38].

It is likely that many countries would aspire to the conditions in Australia as a target for treatment as prevention, as this is a real world population with high coverage of effective treatment. However, new HIV diagnoses, which can be interpreted as reflecting HIV incidence [39], have increased from around 700 cases in 1999 (a nadir of national diagnoses) to around 1,000 new cases annually [33]. This suggests that implementation of treatment as prevention may have less impact on reducing population incidence than previously expected.

Limits to Treatment as Prevention

Treatment as prevention possibly has the greatest chance to succeed now in resource-rich countries with concentrated HIV epidemics, where there is generally universal access to ART, adequate infrastructure, and guidelines that enable early initiation of treatment. However, it is in these very settings that HIV incidence, or surrogate markers thereof, has been increasing [40–45], as in Australia and France. Indeed, at the latest Annecy Group meeting (consisting of representatives from national HIV/AIDS surveillance organisations from developed countries in North America, Australia, Western Europe, and the UK) in Rome in January 2011, it was ascertained that despite differences in epidemiological profiles, surveillance systems, and programmatic responses, HIV epidemics among MSM were generally stable or increasing in almost all of these developed-country settings, despite widespread and increasing availability and effectiveness of ART. Outbreaks of HIV among PWID have also recently been observed in some of these countries [46,47].

There are numerous possible reasons for the apparent ineffectiveness of increased treatment in reducing HIV incidence in "natural experiments". One potential explanation is changes in risk behaviour (shifts in cultural practices, condom fatigue, or risk compensation), as observed in numerous settings including France and Australia [48]. Another possible explanation is the influence of migration from higher prevalence regions, which leads to greater numbers of detected cases in the country of question—sometimes

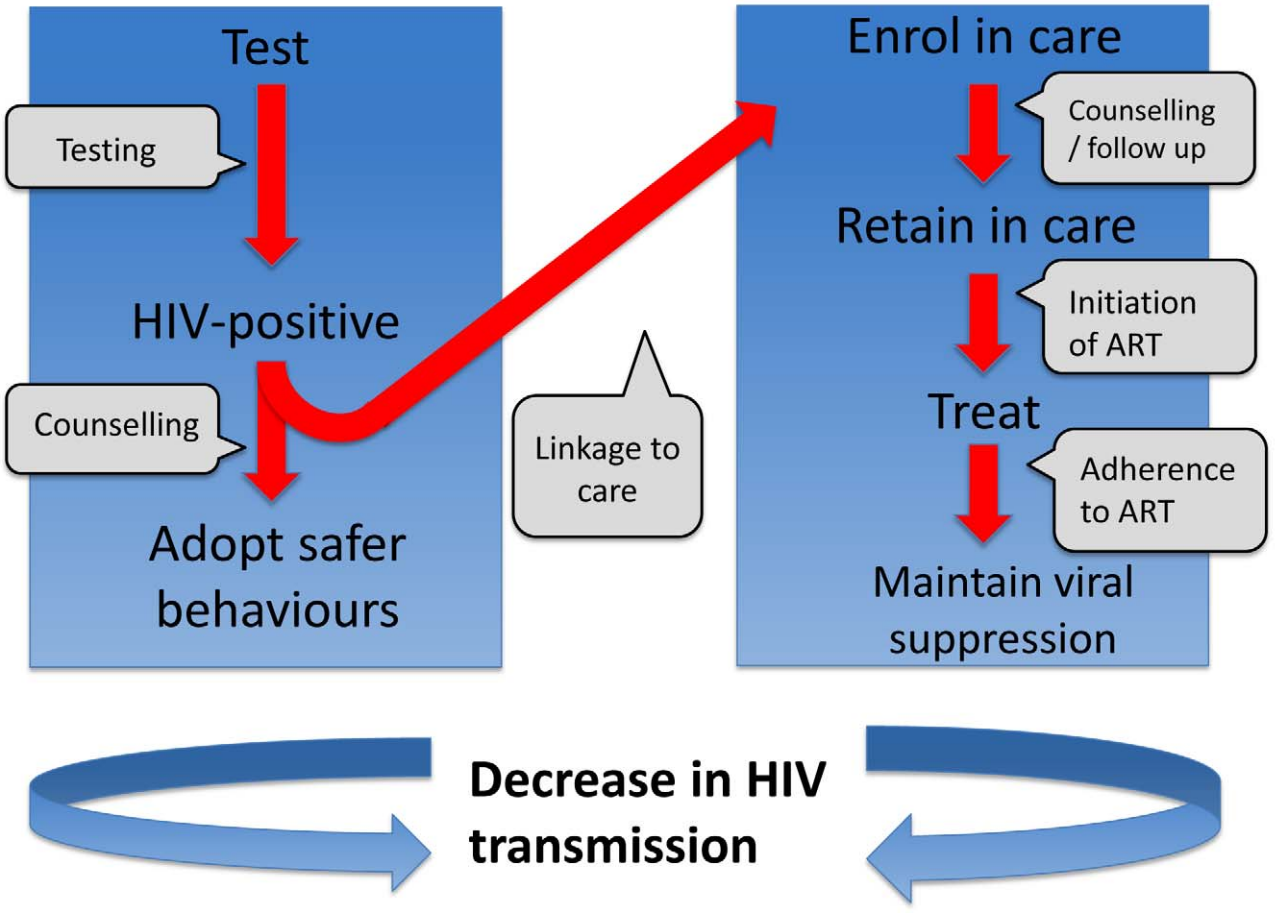

Figure 1. Series of steps required in order to reduce onward transmission from someone infected with HIV.
doi:10.1371/journal.pmed.1001231.g001

used as a measure of incidence—as well as greater background prevalence. There is also the potential emergence of marginalised groups that experience additional barriers to accessing services. These marginalised groups often include migrant and other populations that experience relatively high levels of stigma and discrimination but that are also at greatest risk of HIV infection. Increases in prevalence of other sexually transmitted diseases can also increase HIV incidence, since some sexually transmitted diseases act as a biological cofactor for increasing both HIV infectiousness and susceptibility [49,50]. Another possibility is that treatment is not as effective in reducing infectiousness for riskier modes of transmission as it is for heterosexual transmission (the only mode of transmission considered in the HPTN 052 study) [4,51–53]. Currently, there is little evidence that treatment as prevention is as effective for MSM and PWID [54,55]. These factors may help explain the observed increases in HIV incidence in the era of expanded ART.

One way to consider the problem is that there is a series of barriers to overcome for treatment to be effective in reducing infectiousness (Figure 1). As indicated by Gardner et al. [56], treatment can have a population-level effect in prevention if a high proportion of all HIV-infected people (i) are tested for HIV, (ii) are linked to clinical care in a timely manner, (iii) are retained in care, (iv) receive effective ART, and (v) are adherent to treatment and regularly monitored. It is not uncommon for people to drop out at any of these barriers. Idealised conditions for a treatment-as-prevention strategy may involve setting targets of 90% of all people at each barrier progressing to the next stage. However, as pointed out by Gardner et al., this would result in a maximum of just 66% of HIV-infected people in the population having suppressed virus. Populations of people on ART may have reduced transmission potential, but transmission events are still likely to occur from individuals on ART, as well as from the remaining HIV-infected population without suppressed virus [57].

A related problem of treatment as prevention is that the significant advances in the effectiveness of ART in reducing viral replication have decreased HIV/AIDS-associated mortality [58,59], thereby resulting in a growing pool of HIV-infected people. There is a balance between ART reducing infectiousness and increasing prevalence. This is demonstrated in the natural experiment case of Australia. As shown in Figure 2, the estimated average number of onward HIV infections resulting from each HIV-infected person per year has decreased substantially, but has levelled off at a value above zero. At the same time, the prevalence of HIV has been steadily increasing in Australia because of increased survival due to effective ART (the trend is not altered considerably when adjusted for population size). Correspondingly, overall population incidence has increased over this period. Also, acute HIV infection, with high viremia and high infectiousness, is likely to be an important contributing factor to ongoing transmission [60–63], particularly as most of these cases are usually unrecognised.

On the positive side, the potential problem of there being an increased pool of potential transmitters produced by successful

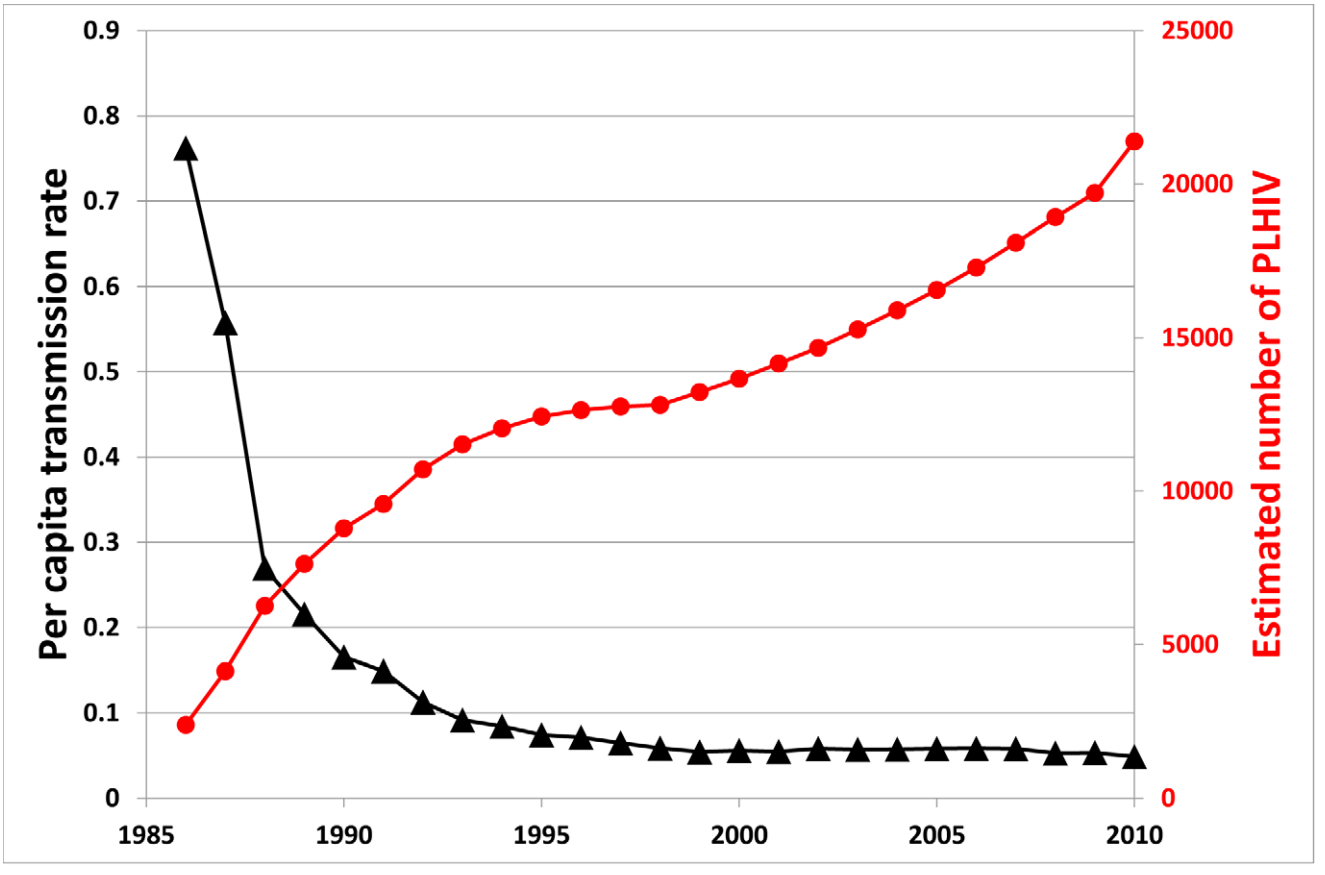

Figure 2. Estimated number of people living with HIV in Australia and per capita transmission rate over time. The per capita transmission rate is defined as the average number of new onward HIV infections resulting from each HIV-infected person per year; this is calculated as the number of new diagnoses in a given year (as a surrogate marker for incidence) divided by the estimated number of people living with HIV (PLHIV).
doi:10.1371/journal.pmed.1001231.g002

ART may turn out to be minor. People's sexual-transmission-related behaviours generally decrease as they age. Therefore, the average transmission rate per infected person may not be reflective of transmission from the majority of people living with HIV. It may be possible to reduce the average reproduction number, R, to below the elimination threshold, $R<1$, even with higher prevalence. Future studies may be able to assess whether this is feasible under realistic conditions.

It is important to note that over the last 5–10 years there have been substantial increases in ART programs across low- and middle-income countries. There is now clear evidence of decreasing HIV prevalence across eastern and southern Africa, which is undoubtedly multifactorial but may reflect some impact of ART on transmission (as assessed by Johnson et al. [64] for South Africa). There have also been large reductions in mortality and corresponding reductions in prevalence throughout Asia associated with ART programs (e.g., [65–67]). Although such benefits are to be celebrated, and there is little doubt that ART programs have likely had an impact in reducing incidence, the levels of undiagnosed infections and treatment coverage make it unlikely that treatment as prevention can lead towards elimination at this stage.

Conclusions

The efficacy of treatment in reducing transmission has been demonstrated for heterosexual transmission in the HPTN 052 trial, with supporting evidence from other types of studies. However, this does not imply that increased ART coverage will result in substantial declines in incidence in real world populations. The average per person rate of transmission will decrease because of ART, but it will likely saturate at a level above zero. Due to increased prevalence of potential transmitters, and other limitations, it may be difficult to decrease overall population incidence

Key Points

- The real world effectiveness of treatment as prevention is likely to be less than the efficacy measured in trials or calculated in optimistic model scenarios.
- Some settings have attained rates of testing and effective ART coverage that many countries would aspire to as targets for treatment-as-prevention strategies.
- Examination of data from treatment-as-prevention "natural experiments" suggests that there are limitations to reductions in population incidence.
- Limitations might stem from behaviour changes, difficulties linking patients with and retaining them in clinical care, differences in the effectiveness of ART for different modes of HIV transmission, and the increasing pool of potential transmitters produced by successful ART.

without other prevention approaches. While trial results are obtained under optimised conditions, where regular counselling and condoms are provided and where there are relatively low rates of sexually transmitted infections, this is often not the situation in the real world. There are also other external factors that may limit the impact of treatment as prevention, including adherence to treatment and shifts in sexual behaviours.

Justifiably, there is large enthusiasm for treatment as prevention. But current planning is based on expected outcomes informed by clinical trials and models—with supporting evidence from ecological and observational studies—that may be overly optimistic. Natural experiments suggest that there are limitations to the overall benefit that can be achieved with this strategy. Before large portions of HIV/AIDS budgets [68] are shifted to treatment as prevention in place of traditional prevention approaches, these limitations need to be given appropriate consideration. It must be acknowledged that ART is cost-effective with respect to clinical benefits [69,70] and is likely to be even more so if prevention benefits are included. But combination prevention using other approaches proven to be effective, feasible, and cost-effective is also essential to reduce incidence among all major groups at risk of infection.

Author Contributions

Performed the experiments: DPW. Analyzed the data: DPW. Wrote the first draft of the manuscript: DPW. ICMJE criteria for authorship read and met: DPW. Agree with manuscript results and conclusions: DPW. Conceived the idea and wrote the manuscript: DPW.

References

1. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
2. Montaner JS, Hogg R, Wood E, Kerr T, Tyndall M, et al. (2006) The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 368: 531–536.
3. Montaner JS, Lima VD, Barrios R, Yip B, Wood E, et al. (2010) Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 376: 532–539.
4. Das M, Chu PL, Santos GM, Scheer S, Vittinghoff E, et al. (2010) Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. PLoS ONE 5: e11068. doi:10.1371/journal.pone.0011068
5. Middelkoop K, Bekker LG, Myer L, Johnson LF, Kloos M, et al. (2011) Antiretroviral therapy and TB notification rates in a high HIV prevalence South African community. J Acquir Immune Defic Syndr 56: 263–269.
6. Anglemyer A, Rutherford GW, Baggaley RC, Egger M, Siegfried N (2011) Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. Cochrane Database Syst Rev 2011: CD009153.
7. Attia S, Egger M, Muller M, Zwahlen M, Low N (2009) Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. AIDS 23: 1397–1404.
8. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. N Engl J Med 342: 921–929.
9. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
10. Holtgrave DR (2010) Is the elimination of HIV infection within reach in the United States? Lessons from an epidemiologic transmission model. Public Health Rep 125: 372–376.
11. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
12. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469.
13. British Columbia Centre for Excellence in HIV/AIDS (2011) Therapeutic guidelines: antiretroviral treatment (ARV) of adult HIV infection—as of January 2011. Available: http://www.cfenet.ubc.ca/sites/default/files/uploads/BCCfE%20Adult%20Therapeutic%20Guidelines_Jan2011.pdf. Accessed 7 December 2011.
14. Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS (2012) HIV treatment as prevention: the utility and limitations of ecological observation. PLoS Med 9: e1001260. doi:10.1371/journal.pmed.1001260
15. Kerr T, Small W, Buchner C, Zhang R, Li K, et al. (2010) Syringe sharing and HIV incidence among injection drug users and increased access to sterile syringes. Am J Public Health 100: 1449–1453.
16. San Francisco Department of Public Health HIV/AIDS Statistics and Epidemiology Section (2009) HIV/AIDS epidemiology annual report 2008. San Francisco: San Francisco Department of Public Health.
17. Scheer S, Kellogg T, Klausner JD, Schwarcz S, Colfax G, et al. (2008) HIV is hyperendemic among men who have sex with men in San Francisco. Sex Transm Infect 84: 493–498.
18. San Francisco Department of Public Health HIV/AIDS Statistics, Epidemiology, and Intervention Research Section (2005) HIV/AIDS epidemiology annual report: 2005. Available: http://sfhiv.org/files/data_reports/hiv_aids_annual_rpt/HIVAIDAnnlRpt2005.pdf. Accessed 10 May 2012.
19. Bourdillon F, Cadore B, Dozon J-P, Gaudin P, Kapusta-Palmer C, et al. (2009 April 9) Statement followed by recommendations on the appropriateness of treatment as an innovative tool for fighting the epidemic of HIV infections. Paris: Conseil National du SIDA.
20. Joint United Nations Programme on HIV/AIDS (2010) Epidemiological factsheet: France. Geneva: Joint United Nations Programme on HIV/AIDS.
21. Haute Autorité de Santé (2009) HIV infection screening in France: screening strategies. Saint-Denis La Plaine (France): Haute Autorité de Santé.
22. Wilson DP, Hoare A, Regan DG, Wand H, Law M (2008) Mathematical models to investigate recent trends in HIV notifications among men who have sex with men in Australia. Sydney: National Centre in HIV Epidemiology and Clinical Research.
23. Jin F, Prestage GP, McDonald A, Ramacciotti T, Imrie JC, et al. (2008) Trend in HIV incidence in a cohort of homosexual men in Sydney: data from the Health in Men Study. Sex Health 5: 109–112.
24. Glynn MK, Rhodes P (2005) Estimated HIV prevalence in the United States at the end of 2003 [abstract].National HIV Prevention Conference; Atlanta, Georgia; 12–15 June 2005.
25. Branson B (2007) Current HIV epidemiology and revised recommendations for HIV testing in health-care settings. J Med Virol 79 (Suppl 1): S6–S10.
26. Branson BM, Handsfield HH, Lampe MA, Janssen RS, Taylor AW, et al. (2006) Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep 55: 1–17.
27. Stekler JD, Golden MR (2009) Learning from the missed opportunities for HIV testing. Sex Transm Infect 85: 2–3.
28. Dodds JP, Johnson AM, Parry JV, Mercey DE (2007) A tale of three cities: persisting high HIV prevalence, risk behaviour and undiagnosed infection in community samples of men who have sex with men. Sex Transm Infect 83: 392–396.
29. Latthaphasavang V, Bouldouyre MA, Rachline A, Ponscarme D, Rozenbaum W, et al. (2012) Antiretroviral therapy initiation in france: adherence to national guidelines and outcome. J Int Assoc Physicians AIDS Care (Chic) 11: 40–46.
30. Le Vu S, Le Strat Y, Barin F, PSurveillance de la syphilis en France, 2000–2006: recrudescence des diagnostics en 2006. illonel J, Cazein F, et al. (2010) Population-based HIV-1 incidence in France, 2003–08: a modelling analysis. Lancet Infect Dis 10: 682–687.
31. Velter A, Bouyssou-Michel A, de Busscher PO, Jauff ret-Roustide M, Semaille C (2007) Enquête presse gay 2004. Saint-Maurice: Institut de Veille Sanitaire.
32. Bouyssou Michel A, Gallay A, Janier M, Dupin N, Halioua B, et al. (2008) Surveillance de la syphilis en France, 2000–2006: recrudescence des diagnostics en 2006. Bull Epidemiol Hebdo 5–6: 39–42.
33. The Kirby Institute (2011) HIV, viral hepatitis and sexually transmissible infections in Australia: annual surveillance report 2011. Sydney: University of New South Wales.
34. Fogarty A, Mao L, Zablotska I, Salter M, Prestage G, et al. (2006) 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. Sydney: University of New South Wales, National Centre in HIV Social Research.
35. National Centre in HIV Research (2011) HIV/AIDS, hepatitis and sexually transmissible infections in Australia: annual report of trends in behaviour 2011. Sydney: University of New South Wales.
36. Falster K, Gelgor L, Shaik A, Zablotska I, Prestage G, et al. (2008) Trends in antiretroviral treatment use and treatment response in three Australian states in the first decade of combination antiretroviral treatment. Sex Health 5: 141–154.
37. The Kirby Institute (2011 September) Australian HIV Observational Database: annual report. Sydney: University of New South Wales.
38. Law MG, Woolley I, Templeton DJ, Roth N, Chuah J, et al. (2011) Trends in detectable viral load by calendar year in the Australian HIV observational database. J Int AIDS Soc 14: 10.
39. Mallitt KA, Wilson DP, McDonald A, Wand H (2011) HIV incidence trends vary between jurisdictions in Australia: an extended back-projection analysis of men who have sex with men. Sex Health 9: 138–143.

40. Grulich AE, Kaldor JM (2008) Trends in HIV incidence in homosexual men in developed countries. Sex Health 5: 113–118.
41. Sullivan PS, Hamouda O, Delpech V, Geduld JE, Prejean J, et al. (2009) Reemergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996–2005. Ann Epidemiol 19: 423–431.
42. Morris SR, Little SJ (2011) MSM: resurgent epidemics. Curr Opin HIV AIDS 6: 326–332.
43. Bezemer D, de Wolf F, Boerlijst MC, van Sighem A, Hollingsworth TD, et al. (2008) A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. AIDS 22: 1071–1077.
44. Jansen IA, Geskus RB, Davidovich U, Jurriaans S, Coutinho RA, et al. (2011) Ongoing HIV-1 transmission among men who have sex with men in Amsterdam: a 25-year prospective cohort study. AIDS 25: 493–501.
45. European Centre for Disease Prevention and Control, World Health Organization Regional Office for Europe (2011) HIV/AIDS surveillance in Europe: 2010. Stockholm: European Centre for Disease Prevention and Control. Available: http://ecdc.europa.eu/en/publications/Publications/101129_SUR_HIV_2009.pdf. Accessed 3 May 2012.
46. Paraskevis D, Nikolopoulos G, Tsiara C, Paraskeva D, Antoniadou A, et al. (2011) HIV-1 outbreak among injecting drug users in Greece, 2011: a preliminary report. Euro Surveill 16. Available: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19962. Accessed 3 May 2012.
47. European Monitoring Centre for Drugs and Drug Addiction (2011) Annual report 2011: the state of the drugs problem in Europe. Lisbon: European Monitoring Centre for Drugs and Drug Addiction. Available: http://www.emcdda.europa.eu/attachements.cfm/att_143743_EN_EMCDDA_AR2011_EN.pdf. Accessed 3 May 2012.
48. Zablotska IB, Crawford J, Imrie J, Prestage G, Jin F, et al. (2009) Increases in unprotected anal intercourse with serodiscordant casual partners among HIV-negative gay men in Sydney. AIDS Behav 13: 638–644.
49. Fleming DT, Wasserheit JN (1999) From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex Transm Infect 75: 3–17.
50. Galvin SR, Cohen MS (2004) The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol 2: 33–42.
51. Fisher M, Pao D, Brown AE, Sudarshi D, Gill ON, et al. (2010) Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. AIDS 24: 1739–1747.
52. Porco TC, Martin JN, Page-Shafer KA, Cheng A, Charlebois E, et al. (2004) Decline in HIV infectivity following the introduction of highly active antiretroviral therapy. AIDS 18: 81–88.
53. Jin F, Jansson J, Law M, Prestage GP, Zablotska I, et al. (2010) Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART. AIDS 24: 907–913.
54. Kelley CF, Haaland RE, Patel P, Evans-Strickfaden T, Farshy C, et al. (2011) HIV-1 RNA rectal shedding is reduced in men with low plasma HIV-1 RNA viral loads and is not enhanced by sexually transmitted bacterial infections of the rectum. J Infect Dis 204: 761–767.
55. Wilson DP (2010) Evidence is still required for treatment as prevention for riskier routes of HIV transmission. AIDS 24: 2891–2893.
56. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ (2011) The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis 52: 793–800.
57. Wilson DP, Law MG, Grulich AE, Cooper DA, Kaldor JM (2008) Relation between HIV viral load and infectiousness: a model-based analysis. Lancet 372: 314–320.
58. Hammer SM, Saag MS, Schechter M, Montaner JS, Schooley RT, et al. (2006) Treatment for adult HIV infection: 2006 recommendations of the International AIDS Society-USA panel. JAMA 296: 827–843.
59. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, et al. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 338: 853–860.
60. Hayes RJ, White RG (2011) Role of acute infection in HIV transmission. Lancet 378: 1913–1914; author reply 1914-1915.
61. Williams BG, Granich R, Dye C (2011) Role of acute infection in HIV transmission. Lancet 378: 1913; author reply 1914-1915.
62. Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, et al. (2011) The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet 378: 256–268.
63. Miller WC, Rosenberg NE, Rutstein SE, Powers KA (2010) Role of acute and early HIV infection in the sexual transmission of HIV. Curr Opin HIV AIDS 5: 277–282.
64. Johnson LF, Hallett TB, Rehle TM, Dorrington RE (2012) The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. J R Soc Interface. E-pub ahead of print. doi:10.1098/rsif.2011.0826
65. Zhang F, Dou Z, Ma Y, Zhang Y, Zhao Y, et al. (2011) Effect of earlier initiation of antiretroviral treatment and increased treatment coverage on HIV-related mortality in China: a national observational cohort study. Lancet Infect Dis 11: 516–524.
66. Zhang F, Dou Z, Ma Y, Zhao Y, Liu Z, et al. (2009) Five-year outcomes of the China National Free Antiretroviral Treatment Program. Ann Intern Med 151: 241–251, W-252.
67. Yang CH, Huang YF, Hsiao CF, Yeh YL, Liou HR, et al. (2008) Trends of mortality and causes of death among HIV-infected patients in Taiwan, 1984–2005. HIV Med 9: 535–543.
68. Kaiser Family Foundation (2011) International AIDS assistance from donor governments: commitments & disbursements, 2002–2010. Available: http://facts.kff.org/chart.aspx?ch=946. Accessed 3 May 2012.
69. Freedberg KA, Losina E, Weinstein MC, Paltiel AD, Cohen CJ, et al. (2001) The cost effectiveness of combination antiretroviral therapy for HIV disease. N Engl J Med 344: 824–831.
70. Badri M, Maartens G, Mandalia S, Bekker LG, Penrod JR, et al. (2006) Cost-effectiveness of highly active antiretroviral therapy in South Africa. PLoS Med 3: e4. doi:10.1371/journal.pmed.0030004

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Considerations in the Design, Conduct, and Analysis of Cluster Randomized Controlled Trials of Combination HIV Prevention

Marie-Claude Boily[1]*, Benoît Mâsse[2], Ramzi Alsallaq[3], Nancy S. Padian[4,5], Jeffrey W. Eaton[1], Juan F. Vesga[1], Timothy B. Hallett[1]

1 Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom, **2** CHU Sainte-Justine Research Centre, University of Montreal, Montreal, Quebec, Canada, **3** College of Nursing Global, New York University, New York, New York, United States of America, **4** Office of the US Global AIDS Coordinator, US Department of State, Washington, District of Columbia, United States of America, **5** Department of Epidemiology, University of California, Berkeley, California, United States of America

Abstract: The rigorous evaluation of the impact of combination HIV prevention packages at the population level will be critical for the future of HIV prevention. In this review, we discuss important considerations for the design and interpretation of cluster randomized controlled trials (C-RCTs) of combination prevention interventions. We focus on three large C-RCTs that will start soon and are designed to test the hypothesis that combination prevention packages, including expanded access to antiretroviral therapy, can substantially reduce HIV incidence. Using a general framework to integrate mathematical modelling analysis into the design, conduct, and analysis of C-RCTs will complement traditional statistical analyses and strengthen the evaluation of the interventions. Importantly, even with combination interventions, it may be challenging to substantially reduce HIV incidence over the 2- to 3-y duration of a C-RCT, unless interventions are scaled up rapidly and key populations are reached. Thus, we propose the innovative use of mathematical modelling to conduct interim analyses, when interim HIV incidence data are not available, to allow the ongoing trials to be modified or adapted to reduce the likelihood of inconclusive outcomes. The preplanned, interactive use of mathematical models during C-RCTs will also provide a valuable opportunity to validate and refine model projections.

Rationale for Cluster Randomized Controlled Trials

Significant progress has been achieved in developing, implementing, and scaling- up safe and effective biomedical and behavioural HIV interventions such as promoting condom use, male circumcision (MC), and the use of antiretroviral drugs for treatment and for the prevention of mother-to-child and heterosexual transmission [1]. Other interventions, such as oral or topical pre-exposure prophylaxis, are in the late stages of clinical evaluation [2]. Considered alone, each intervention provides only partial protection or requires high levels of individual adherence. The combination of several prevention interventions could achieve substantial reductions in incidence even if coverage and adherence to each intervention is suboptimal. The combination approach is widely seen as the most promising way to control the HIV epidemic, especially in highly endemic countries [3,4]. However, the potential population-level effectiveness or impact of combination prevention packages is difficult to predict and needs to be rigorously evaluated in real world settings.

The impact of an intervention at the population level can be very different from its observed efficacy in clinical trials for many reasons, including differences in implementation (e.g., speed and quality of scale-up), target population (e.g., universal, or key subpopulations), and in individual-level factors (e.g., adherence, uptake, sexual behaviour disinhibition) [5–7]. In addition, the level of indirect or herd effects on those not receiving the intervention as a result of the decreasing prevalence of infectious individuals over time is not captured in individual-based randomized controlled trials (I-RCTs) and may differ between interventions [5–7]. Cluster randomized controlled trials (C-RCTs; also called community-based RCTs) are trials in which whole communities, or clusters of individuals, are randomly allocated to receive either the intervention or the control condition [5,8]. C-RCTs can be used to measure the population-level impact of an intervention [5,8]. Typically, the intervention is implemented across the trial communities, but the population-level impact is assessed by measuring the incidence rate among a cohort of individuals in the intervention group compared with a cohort in the control group.

Three large C-RCTs commissioned by the US President's Emergency Plan for AIDS Relief (PEPFAR) to measure the impact of combination prevention packages (including expanded access to

Citation: Boily M-C, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV Treatment as Prevention: Considerations in the Design, Conduct, and Analysis of Cluster Randomized Controlled Trials of Combination HIV Prevention. PLoS Med 9(7): e1001250. doi:10.1371/journal.pmed.1001250

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: TBH thanks the Bill & Melinda Gates Foundation and The Wellcome Trust for funding support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mc.boily@imperial.ac.uk

Abbreviations: ART, antiretroviral therapy; C-RCT, cluster randomized controlled trial; I-RCT, individual-based randomized controlled trial; JHU/USAID, Johns Hopkins University/United States Agency for International Development; MC, male circumcision; PEPFAR, US President's Emergency Plan for AIDS Relief

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

antiretroviral therapy [ART]) on HIV incidence in different populations will start shortly (Table 1) [9–11]. The different trial intervention packages focus on the scale-up of ART (i) initiated at different CD4 levels in Zambia and South Africa, (ii) prioritising those with the highest viral loads in Botswana, and (iii) in combination with other interventions in Tanzania.

In a context in which resources generally are becoming increasingly scarce, obtaining valid answers from these trials will be critical for the future of HIV prevention. Positive results showing large reductions in HIV incidence could shift the paradigm guiding the response to HIV epidemics, whilst negative results could challenge the case for continued investment in combination prevention interventions.

Despite being considered the gold standard for measuring the population-level impact of interventions, the design, implementation, and interpretation of C-RCTs can be extremely challenging [5,8,12,13]. In the past, some researchers have turned to mathematical models once the studies were completed to help understand ambiguous and counter-intuitive results from C-RCTs [14–16]. Others have advocated for their use before studies begin to improve trial design [5,17–21]. All three PEPFAR trials currently include an HIV transmission dynamic modelling component to complement traditional statistical approaches for the analysis of C-RCTs. Mathematical models will be used in three distinct phases—at the formative stage of trial planning, during the trial itself to monitor progress, and at the end of the

Table 1. Main characteristics of cluster randomized controlled trials for combination prevention of HIV transmission commissioned by PEPFAR.

Study	CDC/HSPH[a]	JHU/USAID	PopART (HPTN 071)
Site	Botswana	Iringa, Tanzania	Zambia+South Africa (Western Cape)
Number of arms	2	2	3
Intervention arm(s)	A: Enhanced HIV testing (including mobile and home-based testing), active linkage to care and treatment; enhanced MC; ART for all HIV-infected persons with CD4<350 cells/μl or with HIV-1 RNA>10,000 copies/ml; and point-of-care CD4 testing in antenatal clinics with universal HAART in pregnancy started by 28 wk gestation, as well as HIV retesting at delivery among women HIV-negative in second trimester or earlier	A: Treatment by CD4<350 cells/μl; active scale-up and linkage to MC; cash transfer for young women; targeted outreach to the most at-risk populations (including female sex workers); social and behaviour change communication	A: Universal community home-based testing; active linkage of HIV-positive individuals to care and immediate ART according to national guidelines and/or MC. B: Same as A but ART at CD4<350 cells/μl
Control arm	B: Standard of care[b]	B: Standard of care[c]	C: Enhanced standard of care[d]
Design	Pair matched	Stratified	Triplet matched
Number of randomized clusters			
Total	30	24	24 (South Africa: 9, Zambia: 15)
Per arm	15	12	8
Average size of randomized cluster	5,800	8,000–10,000 (~55%>15 y)	50,000 (25,000>18 y)
Overall cohort followed up			
Age eligibility	16–64 y	15–39 y	18–44 y
Size per cluster	~500 adults per cluster	~500 adults per cluster	~2,500 adults per cluster
Total size	15,000	12,000	60,000
Primary Outcome	HIV incidence[e]	HIV incidence[e]	HIV incidence[e]
Follow-up duration	3–4 y	2 y	2 y
HIV incidence assumption	~1.5 per 100 person-years	1 .0–1.5 per 100 person-years	1.0–1.5 per 100 person-years
Anticipated HIV prevalence at baseline	25%	10%–15%	15%
Target reduction in incidence	In arm A versus B: ~50%	In arm A versus B: ~40% (35%–50%)	In arm A versus C: −50% to 60%; in arm B versus C: −25% to 30%
Stages when modelling is currently planned	Start	Start, interim, final	Start, final
Status	Planning	Pre-trial	Pre-trial

Data as of 15 March 2012.
[a]The design of the intervention and plan of analysis for this trial are still being finalised.
[b]Standard of care is ART for HIV-positive individuals with CD4<350 cells/μl or AIDS.
[c]Standard of care is standard referral to MC and ART according to Tanzania guidelines (this will soon change from CD4<200 cells/μl to CD4<350 cells/μl, initially focusing on HIV-positive people with tuberculosis and pregnant women).
[d]Standard of care is no home-based testing or home-based visit to facilitate linkage to ART. ART given according to country guidelines; standard referral to MC.
[e]Cumulative HIV incidence measured over the trial duration.
CDC/HSPH, US Centers for Disease Control and Prevention/Harvard School of Public Health; HAART, highly active ART; JHU/USAID, Johns Hopkins University/United States Agency of International Development; PopART (HPTN 071), HIV Prevention Trials Network.
doi:10.1371/journal.pmed.1001250.t001

trial to assist in interpretation and evaluation of short- and long-term impact.

In this review, we draw on results from a range of models to identify important considerations that should inform the design and interpretation of C-RCTs of combination interventions. We then propose how mathematical modelling can be integrated into the design, conduct, and analysis of the planned trials to complement traditional statistical approaches.

Considerations for the Design, Conduct, and Interpretation of Cluster Randomized Controlled Trials

Previous modelling studies suggest that ART used alone or in combination with other interventions could significantly reduce long-term HIV transmission [4,10,22–26]. However, to evaluate the impact of interventions in the time frame of a trial, which is usually 2–3 y, it is critical to understand what magnitude of impact can be expected in the short term, whether the short-term impact is predictive of the long-term impact, and what implementation efforts might be required to achieve the desired level of impact. The answers to these questions are influenced by different determinants of the magnitude of intervention impact, and of the measurement and assessment of impact in C-RCTs. The important considerations and implications for C-RCTs for these determinants are summarised in Table 2. We provide illustrations of the main points below.

Determinants of the Magnitude of Intervention Impact

Increase of intervention impact over time. A concern of particular relevance for C-RCTs is that the full impact of interventions on HIV incidence at the population level is unlikely to be generated immediately after the start of the trial [16,26]. For example, HIV risk might actually increase during the wound healing period following MC procedures [27]. In the case of ART, complete viral suppression and reduced infectivity takes time to occur after initiating treatment. Moreover, if ART eligibility is not immediate but occurs only once an individual reaches a predetermined CD4 level, as shown in Figure 1, there will be a lag between the start of the screening and treatment programme and the time point when the fraction of eligible HIV-positive individuals provided with ART is large enough to reduce transmission at the population level. This differs from I-RCTs, in which all eligible patients in the trial are immediately provided with their assigned treatment. In addition, in real-life situations, ART failure, poorer treatment adherence, and viral blips may be more frequent than in the ideal conditions of trials such as HPTN 052 [28], thereby reducing intervention impact. Finally, indirect benefits or "herd effects" accrued through the prevention of onward transmission, which are measurable in C-RCTs but not in I-RCTs, manifest more slowly, as these rely on a decreasing prevalence of HIV infections in some subpopulations [5–7].

Thus, C-RCTs designed to evaluate intervention impact after a short time will assess an impact that has not reached its maximum potential [16,26]. For example, in Figure 1, HIV incidence is reduced by only 34% at 2 y even with a very ambitious combination intervention, compared with 66% after 25 y (not shown). Studies that estimate the intervention impact from changes in HIV prevalence, as is commonly done when monitoring key populations, have an even slower increase in intervention impact [13,29]. Finally, because it can take different amounts of time for each intervention component to have its full effect, the overall impact of a combination intervention may be most strongly determined by different components at different time points after the start of the intervention programme (Figure 1) [26].

Influence of the epidemiological context. The epidemiological context for a given country or population is determined by the drivers of HIV transmission (e.g., patterns of risk behaviour and contact, and key biological factors that facilitate transmission) and by the past trajectory of the epidemic, which determines the distributions of individuals at different stages of HIV infection [30–36]. The underlying patterns and strength of transmission interact with the intervention and make predictions more complex. For example, for interventions that include expanded access to ART to prevent HIV (as will be the case in the three trials summarised in Table 1), the amount of transmission by an individual before treatment initiation, including during the initial highly infectious period, will determine the level of treatment required to reduce incidence. The amount of transmission generated early after infection depends on the number of concurrent sexual partners, the interval between sexual partnerships, the frequency and type of sexual acts, transmission probabilities, the fraction of new sexual partners who are already infected, and the prevalence of cofactors of HIV transmission, such as other sexually transmitted infections [36–39].

The effect of the same universal "test and treat" intervention can differ greatly across populations that have similar HIV prevalence, incidence, and rate of partner change but differences in other key sexual behaviours [31]. For instance, an intervention may reduce incidence by nearly 100% and eliminate the infection in one population if there is little heterogeneity in risk behaviour, whereas exactly the same intervention may reduce incidence by only 60% in another population if there is substantial heterogeneity and assortative mixing by sexual activity levels [31]. In a heterogeneous population transmission can persist within the highest risk group because individuals transmit rapidly after becoming infected and before getting ART. Thus, the impact of the same intervention may vary across C-RCTs conducted in different populations or settings, and, consequently, the findings from one trial may not necessarily apply to another setting. Mathematical models can take into account knowledge of the drivers of the HIV epidemic and the intervention impact in a specific trial setting, and help generalise trial results to other epidemiological contexts [5,13,21,40].

Identifying drivers of short-term and long-term intervention impact. Although C-RCTs aim to measure the impact of interventions over a short period, broader public health interests are usually longer term. Factors that drive short-term impact may not be the same as those determining long-term impact and overall success of the programme. For example, one would expect the short-term impact of ART for prevention to be driven by factors such as the speed of linkage and retention in care during the first years after treatment initiation and adherence in the months following initiation, whereas long-term impact would be more sensitive to factors such as prolonged maintenance of retention in care and high adherence, continued frequent HIV testing, and robust linkage to care [22,23,26,31]. Collection of data on these long-term factors may not be immediately useful for understanding the trial results in the short term, but will help predict the long-term impact of the trial results.

Finally, one important and often neglected consideration for C-RCTs is that most modelling analysis assumes that the intervention coverage is uniform with respect to different forms of risks and geography. This is unlikely to be the case in real world settings, as it is difficult to rollout an intervention with equal intensity in all settings, particularly if accessibility and outreach to key populations is poor [4,22–26]. Modelling of a C-RCT of mass treatment of sexually transmitted diseases in Rakai, Uganda, showed that

Table 2. Summary of important considerations for the design and interpretation of cluster randomized controlled trials (of combination interventions.

Important Considerations	Implications for Trials
Determinants of the magnitude of intervention impact	
Increase in intervention impact following the start of trial can be slow due to a number of delays before the full impact develops	Short-term impact will underestimate the long-term impact; substantially reducing HIV incidence over a trial of short duration will be challenging even with an ambitious combination intervention and rapid scale-up; it is important to set realistic expectations about the achievable magnitude of impact over the trial duration; this slow growth in impact can undermine the utility of stepped-wedge designs (with staggered randomized time of delivery of the intervention in each community) to measure a difference in HIV incidence between different interventions or components because the time interval between steps may need to be unfeasibly long[a]
The maximum impact of different intervention components is achieved at different times	The trial duration will influence which type of intervention seems to be the most effective; the overall impact of a combination intervention will be most strongly determined by different components at different times
The epidemiological context influences the intervention impact	The impact of the same intervention may not be the same across trials conducted in different epidemiological contexts; the results of the trial may not be generalisable to other settings
HIV prevalence and HIV incidence do not exhaustively describe the epidemiological context	This may introduce imbalance between the intervention and control arms, even after matching for HIV prevalence or even HIV incidence
The drivers of short-term and long-term impact can be different	Sufficient information on the epidemic drivers should be collected during the trial to help interpret trial results and to predict longer term impact
Distribution of coverage matters even at high coverage	Intervention impact can be substantially reduced if the intervention does not reach high-risk individuals; intervention impact can be substantially improved if the intervention does reach high transmitters; to understand trial results, detailed information on programmatic (e.g., coverage, uptake) and intermediate outcomes (e.g., change in behaviour, CD4 levels, viral load) by risk groups, age, and clinical status in both the intervention and control communities will be essential
Challenges to the measurement of impact	
Measurement of HIV incidence in a cohort over the whole trial duration, before the intervention has reached its full effect, underestimates the change in incidence that is achieved at the end of the trial	It would be better to measure incidence at the start and end of the trial using two independent cohorts with shorter follow-up
Evolving standard of care in control arm, as the coverage or scale-up of standard of care may improve over time	Reduces the contrast between intervention and control communities over time; our ability to measure a difference between trial arms will depend on the rapid scale-up of the intervention, having a large number of clusters to enable detection of smaller effects, or having trial duration longer than 2–3 y, to allow the intervention impact to be seen
Imbalance in key epidemiological characteristics between trial arms can occur, as HIV incidence and prevalence do not determine all key epidemiological characteristics that influence intervention impact	Could lead to a spurious indication that the intervention is working better or worse than it really did—matching clusters may be desirable; matching on HIV prevalence alone may not be sufficient, as trajectories in incidence and underlying patterns of risk behaviour across trial communities would not be captured
Dilution and contamination of the intervention impact may occur due to movement and sexual partnerships across multiple communities	The influence of the different sources of contamination on trial results will depend on the type of intervention; when there is extensive sexual contact between individuals from the trial arms, the measurable impact may be more strongly determined by acquisition-reducing than infectiousness-reducing interventions, such as ART; choosing distinct, independent communities will be important, especially to evaluate ART interventions

[a]Stepped-wedge design can still be useful for programme and intermediate outcomes, as changes in these outcomes can occur more rapidly than for HIV incidence or prevalence.

doi:10.1371/journal.pmed.1001250.t002

even if a high coverage is achieved overall, differential coverage in which those with highest sexual activity are not reached can severely attenuate the impact of the intervention [15]. Conversely, if those at highest risk can be effectively prioritised as coverage is increased, the impact of interventions can be enhanced [15,32,40]. Thus, collecting detailed information on programmatic, implementation, and intermediate outcomes (e.g., changes in behaviour, CD4 levels, and viral load) by risk group, age, and clinical status in both the intervention and control communities at different times during the trial is necessary for evaluation of the short-term and long-term impact.

Challenges in Measuring Impact

Even if the intervention really does have an impact following rapid scale-up, high uptake, good adherence, etc., external factors may compromise our ability to measure a difference in impact between the intervention and control clusters.

Measuring HIV incidence over the whole trial duration. When incidence is measured in a single cohort over the whole duration of a trial, as currently planned in the three combination intervention trials (Table 1), the measured difference in incidence between the trial arms will be attenuated compared with the true difference that would be seen if HIV incidence were measured only at the end of the trial (Figure 2) [16]. This is because the measurement of incidence includes exposure while the intervention activities are still being ramped up and have not yet reached their full impact. Ideally, incidence should be measured at the start and end of the trial, using two independent samplings of the cohorts with shorter follow-ups. However, this solution may not be feasible in practice because of time constraints or costs. Thus, caution must be used when using modelling predictions of intervention impact based on predicted incidence at fixed time points (i.e., an instantaneous reduction in incidence) to estimate effect size and inform trial design.

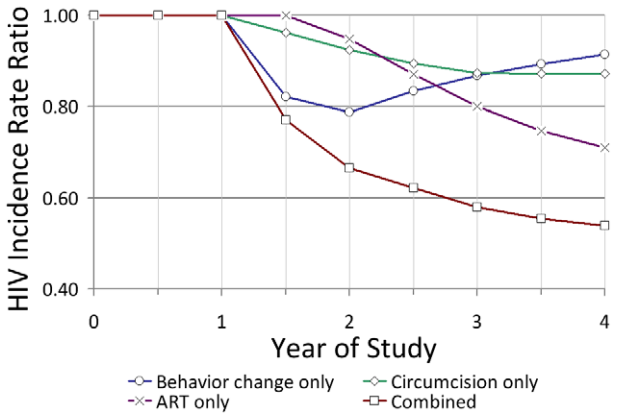

Figure 1. Predicted short-term impact of three intervention components linked to HIV testing in KwaZulu-Natal, South Africa. The model is based on a high-transmission setting under conditions of the current standard of care versus a high-coverage combination intervention (see [26]). The instantaneous HIV incidence rate ratio in the *y*-axis is intervention versus control. Impact estimates include an initial 6-mo period of preparation for the study. Assumptions for the combination intervention: 90% of adults in the intervention community are tested in the first year and thereafter every 4 y; those who test positive reduce risk behaviour for 3 y (on average) (25.0%/12.5% of men/women increase condom use; 25%/25% reduce partner acquisition); 70% of uncircumcised men are circumcised in the first year (efficacy = 60%); and all those in need of treatment (CD4 cell count <350 cells/μl) are immediately treated with ART (efficacy = 92%) with an annual dropout rate from treatment of 5%. The efficacy of MC in reducing susceptibility is assumed to be immediate (i.e., the wound healing period is negligible). Viral suppression for infected individuals once on treatment is immediate (i.e., no delay between treatment initiation and viral suppression). Assumptions for the standard of care: 20% of individuals test annually; 12.5%/6.5% of men/women who test positive increase condom use, and 12.5%/12.5% reduce partner acquisition, for one year; HIV-positive individuals are treated if CD4<200 cells/μl (dropout rate of 15%); and 27% of men are circumcised at baseline and 10% more over 4 y since the start of the intervention.

doi:10.1371/journal.pmed.1001250.g001

Evolving standard of care in control arm. One of the strengths of C-RCT design is that it has a control group. One inescapable challenge, especially for the Johns Hopkins University/United States Agency for International Development (JHU/USAID) study, is that coverage with the standard of care in the control arm may increase over time, albeit more slowly than in the intervention arm, because of ongoing scaling-up activities for MC and/or changes in ART guidelines (from CD4<200 cells/μl to CD4<350 cells/μl). This can potentially reduce HIV incidence in the control arm and thereby reduce the contrast with the intervention arm, so compromising the power of the trial.

Imbalance between trial arms. One important and rarely acknowledged implication of the epidemiological context is that it could introduce an imbalance between trial arms, despite randomization and even if clusters are matched according to HIV incidence and/or prevalence. Such imbalance could lead to biases in either direction [8,16]. Measurements of baseline HIV incidence before the start of the trial intervention, allowing the evaluation of "within cluster" changes in HIV incidence (before–after comparison), could help reduce this problem. However, this approach may not necessarily eliminate all confounding if differences in baseline HIV incidence actually reflect differences in key baseline epidemiological characteristics that influence how each community responds to interventions. Statistical adjustment limited to differences in cluster-level prevalence (or incidence) may only partially control for these nonlinear effects, especially if valid measures of most of the key potential confounding factors, and their interactions, are not available. Despite the benefit of randomization, which protects against known and unknown confounding, imbalance remains of particular concern in C-RCTs, as fewer units are randomized than in I-RCTs. For instance, there will be ~24–30 clusters in the three planned C-RCTs versus ~2,000 individuals in many I-RCTs [12].

Ideally, the number of clusters that are randomized needs to be sufficiently large to minimise the risk of imbalance or to allow matching of pairs or triplets of similar clusters, as proposed in the US Centers for Disease Control and Prevention/Harvard School of Public Health (CDC/HSPH) and PopART trials, using

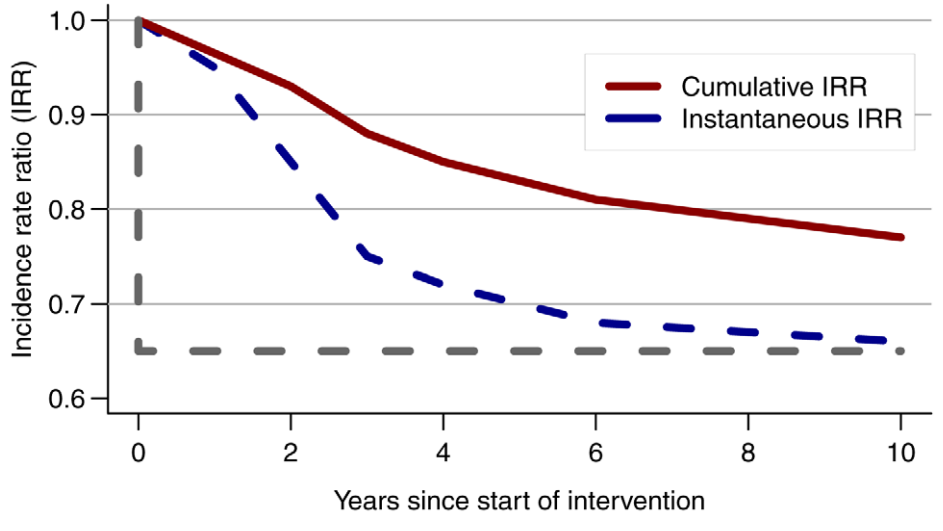

Figure 2. Consequence of measuring HIV incidence over the whole trial duration. Comparison of the instantaneous reduction in HIV incidence measured at one time point with the cumulative incidence rate ratio (IRR) measured over the whole trial duration (i.e., in a cohort that was initiated at the start of the trial) in a simulated population in Zimbabwe [16]. The grey dotted line shows the IRR if the full impact were achieved at the start of the intervention rather than after 10 y. The instantaneous IRR is 0.65 compared with only 0.77 for the cumulative IRR at year 10. From [16].

doi:10.1371/journal.pmed.1001250.g002

information on the epidemiological indicators available at the start of the trial. Whilst matching should help increase power if the matching indicators are highly correlated with the primary outcome [8], it can also be inefficient and reduce power if the matching indicators are not strongly related to outcomes. This could be the case when using only estimates of HIV prevalence. In addition, matching using several factors might not be feasible, as only a limited number of communities are available for most C-RCTs, and this might also limit the types of analyses that can be done [8,41]. Due to limited information, especially on HIV prevalence at the cluster level in Iringa, Tanzania, a stratified approach is being adopted in the JHU/USAID trial.

Dilution and contamination. To minimise the risk of contamination, the clusters enrolled should be distinct, independent epidemiological units. Risk of contamination increases when individuals move or form sexual partnerships between clusters in different intervention arms of the trial or communities not enrolled in the trial. Individuals can also be lost to follow-up or can access an intervention not assigned to their cluster, thereby diluting the differences between arms.

The influence of the different sources of contamination on trial results will also vary for different interventions. For example, the impact of interventions that reduce acquisition of HIV, such as MC, should be only modestly affected by sexual mixing between communities, as long as residents in the intervention community are sufficiently exposed to the intervention. However, if substantial mixing occurs between communities, then interventions that reduce infectiousness such as ART may not have an observable impact in the intervention community. Choosing communities that are more isolated will therefore be more important for evaluating treatment as prevention than behaviour change communication or MC interventions. Although the risk of dilution and contamination can be minimised by choosing geographically separated communities, studies should still aim to collect data on sexual partnerships between communities; genetic sequencing technologies may be a useful for this [8,42].

The Role of Modelling in Planned and Future Cluster Randomized Controlled Trials

As discussed above, mathematical models have been useful to highlight important considerations relevant to C-RCTs. Based on this prior knowledge, we describe how mathematical models can be used before, during, and at the end of trials with reference to the three planned PEPFAR trials (Table 1), and with suggestions for future trials that may be planned subsequently.

Modelling Prior to the Start of the Trial: Formative Phase

Informing design and intervention targets. Prior to the start of the trial, provided that sufficient data are available, models can be used to better understand the epidemic drivers in the trial communities and to define the combination intervention package most suited for the epidemiological context [40]. Then, models can be used, as in the three planned C-RCTs, to estimate the potential impact of the selected intervention in a given setting and to simulate how large a difference in HIV incidence (or prevalence, which is often used for key populations) will develop between the study arms over the trial duration, and how quickly it will develop. These impact estimates should take into account that the prevention activities occurring in the control arm may also evolve over the trial duration [13,43]. Models can also be used to inform the minimum programmatic and implementation targets, such as the speed of scale-up and coverage of each intervention component, and/or the intermediate outcomes, such as change in behaviour, that are required to achieve the desirable impact or "effect size" at the end of the trial. Together, this information contributes to the overall design of the study.

Once a study design is chosen, models can also be used to simulate the process of the trial to identify potential difficulties such as the influence of sources of contamination or imbalance, to evaluate gain in power from matching clusters, or to validate sample size and power calculations [5,16,20]. All three C-RCTs are using models to simulate the influence of possible contamination. In addition, simulations can be used to control the chance of obtaining spurious significant results (type I error) when a novel design, such as an adaptive design that allows preplanned mid-course corrections, is used (see section on interim modelling analyses below) [5,16–18,20,47–53].

Refinement of intervention package. Once calibrated to the specific trial setting using techniques previously described [13,44–46], models can be used to refine the combination intervention package by assessing the impact of the different intervention components, such as promotion of condom use, MC, or ART, independently and in combination. This assessment can be achieved by varying the coverage, intensity, and uptake in different risk groups in the models. These modelling analyses help identify the minimum combined package (in terms of effort, persons reached, and resources spent) needed to maximise the short- and/or long-term impact, since the optimal package may depend on the time frame used to assess it [5,26,27,32,33]. These analyses can provide useful information about the attenuation of impact that could ensue from worse coverage in populations at the highest risk of infection, or from scaling up one component more quickly than another.

Modelling during the Trial: Interim Modelling Analyses

Although statistical methods for formal interim efficacy review of phase III I-RCTs can theoretically be adapted for monitoring C-RCTs [52], they may be logistically more challenging, especially for short C-RCT trials, if HIV incidence measurements are required soon after the start of the trial. We propose the innovative use of mathematical modelling to conduct interim analyses, when interim HIV incidence data are not available, to allow the ongoing trials to be modified or adapted to reduce the likelihood of inconclusive outcomes.

The planned C-RCTs commissioned by PEPFAR are particularly ambitious, aiming to reduce HIV incidence by 25%–60% in just 2 or 3 y (Table 1). As currently proposed by the JHU/USAID team, mathematical modelling can be used to help monitor the progress of the trial. This can help assess the quality of the implementation and, if needed, trigger predetermined mid-course corrections as part of an adaptive design, such as accelerated roll-out or modified trial duration [48–51]. For example, a minimum level of coverage (at specific time points) under which the trial will probably be unsuccessful can be predetermined. In addition, interim modelling analyses can be done using additional data from the baseline surveys in each trial cluster (such as sexual behaviour and updated HIV prevalence estimates) and the most recent information on process indicators of coverage and intensity that is available. Robust monitoring and evaluation data will be necessary to permit these kinds of analyses in a timely fashion. The objective is to predict the likely impact at the end of trial and to estimate the probability that an effect size will be detected. This is similar to a conditional power analysis for futility stopping conducted at the interim analysis of an I-RCT, after which the trial is stopped if the interim results suggest that the effect sought is unlikely to be achieved if the trial continues. This approach is particularly relevant in situations in which no interim incidence measurements are available to conduct a formal interim analysis.

The information gained from this type of modelling can then be used to guide the conduct of the rest of the trial (Figure 3). The question of particular interest is to determine, with the level of coverage and intensity achieved between baseline and interim analysis, the likelihood of observing a measurable impact at the end of the trial and whether changes to the implementation of the intervention or conduct of the study are required to maximise its usefulness. When considering allowing modifications of some prespecified aspect of the design based on interim analysis, it is necessary to consider its possible influence on the overall type I error (chance of detecting a false positive result). Although the type I error is usually well controlled with traditional (non-adaptive) trial designs, this is generally not the case for adaptive trial designs, where inflation of the type I error is often a concern [48,49,53]. Thus, mid-course corrections should be carefully planned and implemented using trial simulations to demonstrate that the type I error will be protected [49–51,53]. The interim modelling analysis may come to one of the four conclusions shown in Figure 3. For example, a finding that there is little chance of detecting an impact even if the study lasted longer (outcome iv) would indicate a high likelihood of obtaining non-informative results, akin to the concept of "futility" in I-RCTs.

This information should be used as a warning of potential problems, and the recommended action might include improving programmatic targets with or without increasing study duration. Those decisions should be discussed within the framework of the independent data monitoring committee that oversees the conduct of the trial, the quality of the implementation, and impact projection. The data monitoring committee could endorse the trial protocol team's decision and/or recommend modifications of the trial. At least one or two members of the data monitoring committee should have expertise in mathematical modelling.

Modelling at the End of the Trial: Evaluation, Interpretation, and Extrapolation

Depending on the outcome of the trial, models can be used in slightly different ways to help interpret the trial results (Figure 4) [5,13–16,29,31,43,44]. The first goal of this final set of analyses is to test and potentially validate final model predictions of intervention impact at the end of the trial. To do this, the analyses should use all the relevant available data on sexual behaviour as well as process indicators of intervention coverage and intensity collected in each community and trial arm during the whole trial duration, to inform prior model parameter distribution and calibrate the model to the HIV outcomes. For validation purposes, model predictions should ideally be derived just before the end of the trial, while the modeller is still blind to the empirical trial results on HIV incidence.

If the model predictions and trial results are similar, then this validates the model projections, and the model can be used for further analyses with a greater degree of confidence. If not, refinements in the statistical analysis, such as adjustment for baseline factors, and/or in the mathematical model are required until the source of the discrepancies is identified, as shown in Figure 4.

If the trial results suggest that the intervention has a significant impact and there is no imbalance in key indicators of epidemi-

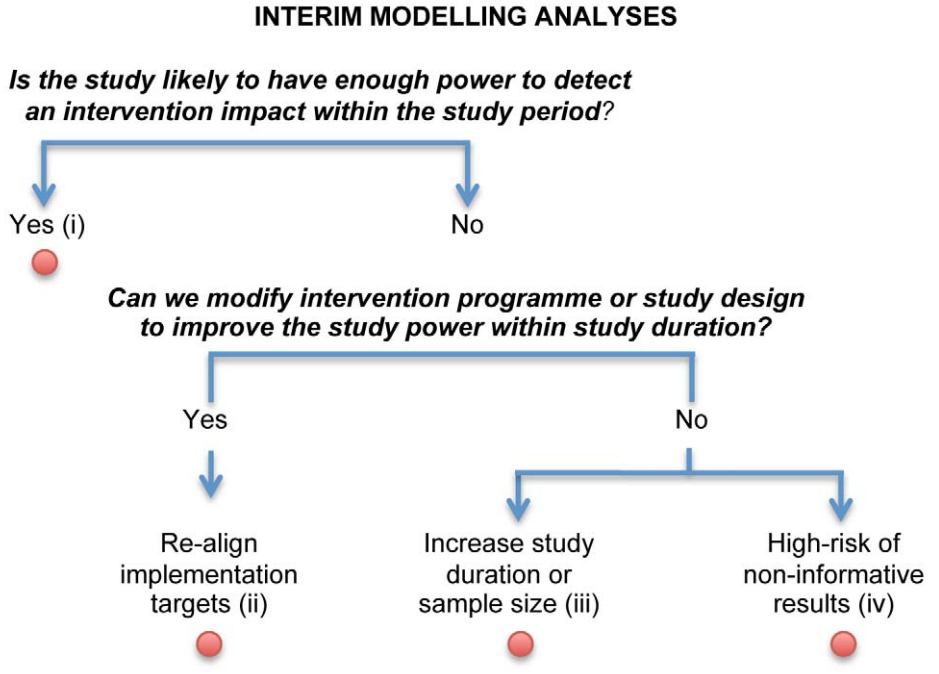

Figure 3. Logical flow of interim modelling analyses. This approach uses available data from the baseline surveys in each trial cluster and information on process indicators of coverage and intensity available for each cluster within each trial arm gathered after the start of the trial. These data would not include observed HIV incidence. The interim modelling analysis may come to one of four conclusions. (i) The targeted effect size on HIV is likely to be achieved at the end of the study without having to modify the intervention targets/implementation strategy. (ii) The targeted effect size is unlikely to be achieved, and therefore the intervention targets/implementation strategy need to be revised. (iii) The targeted effect size is unlikely to be achieved, even if the intervention targets are improved to their realistic maximum, unless there is a change in the study design (such as an increase in sample size or study duration). (iv) There is little chance of being able to detect an impact at the end of the trial even if the study duration is increased. The number of interim analyses should be predetermined at the start of the trial and take into account trial characteristics, logistical considerations (such as the time and cost required to regularly update programmatic data during the trial and to perform the modelling analyses), and the statistical effect of the interim analysis and proposed changes on the overall type I error.
doi:10.1371/journal.pmed.1001250.g003

ological context between the trial arms, the final modelling analysis can predict the number of infections that would be averted by the combined package in the intervention arm compared with the standard package in the control arm over different time periods if the intervention were continued. The counterfactual would be simulated using the level of coverage and change in behaviour and other programmatic outcomes observed with the standard package in the control arm over the course of the study. The models could also be adapted to project impact in other populations with different epidemiological contexts to help generalise the trial results, and to compare the results with those of other trials of combination interventions. Provided that good costing data are also collected (as is planned in all three trials), it will be possible to link the costing data to the short- and long-term model predictions for a cost-effectiveness analysis [54]. One of the challenges will be to understand the costs incurred for the intervention in the trial, including start-up, small-scale set-up, and cost of the learning curve, compared with how these costs would evolve in a large-scale programme over the long term [54].

Currently, C-RCTs are not designed to establish differences between the different intervention components, as this would require larger trials with multiple arms, potentially using factorial designs. It may be possible to model and predict the impact of each specific component of the intervention package independently, but it will be challenging. If individuals were exposed to several intervention components during the trial, it would be difficult to attribute an observed reduction in risk behaviour, e.g., relating to sexual behaviour or adherence, to one particular component. Also, with the acknowledged limitations of the collection of behavioural data, it is difficult to reliably transduce the effects of reported changes in behaviour into an impact on transmission. It may be more feasible to link interventions that have hard end points, such as being circumcised or starting on ART, to estimated impact. The epidemiological synergy between interventions, which can make the impact of combination prevention greater than the sum (or multiplication) of its parts, may also be an important part of the total impact. Conversely, redundancy between components may reduce the combined intervention impact, meaning that the total intervention impact may be lower than the sum (or multiplication) of its parts.

If there is a significant imbalance in key baseline characteristics between trial arms, it would be useful to assess the extent to which this imbalance could have biased the observed impact estimate, and to produce "adjusted" estimates, i.e., estimates revised upward in the case of a positive trial or downward in the case of a negative trial.

Finally, if a trial produces negative results despite the coverage of interventions such as ART and MC increasing substantially, the main points to explore would be the following: to what extent this lack of impact was because the trial was too short; how long would it have taken to detect a measurable impact; and whether the level of contamination in the control group was too high.

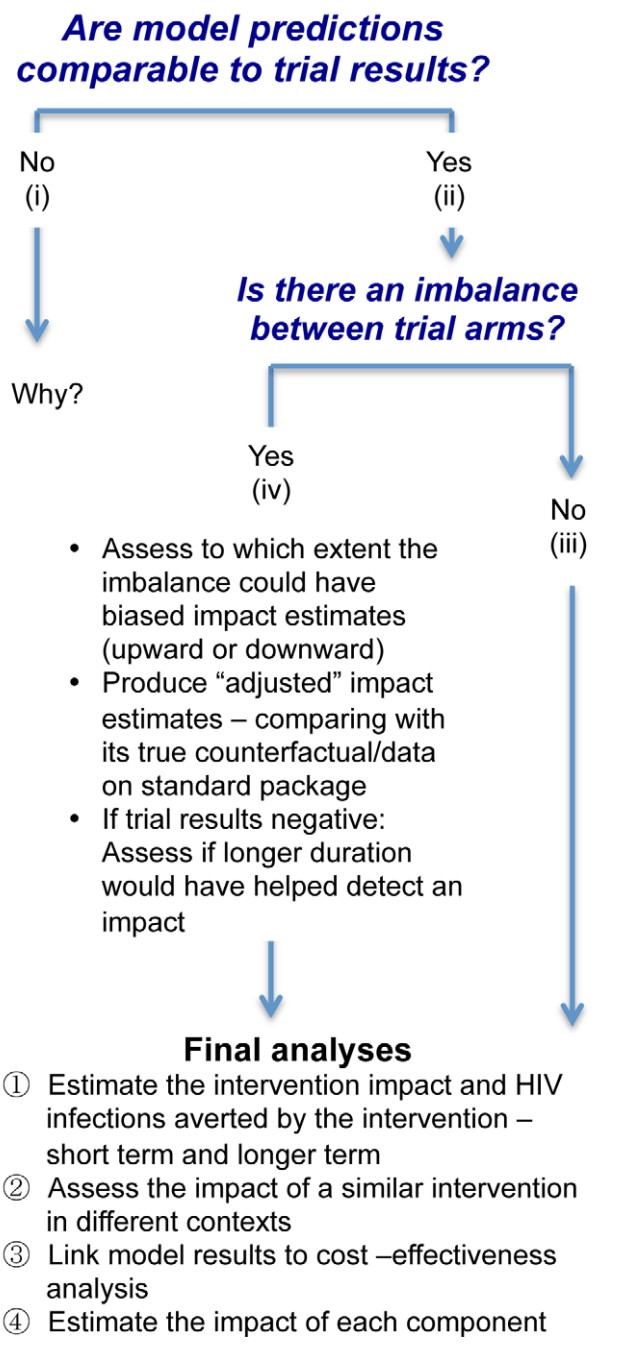

Figure 4. Logical flow of modelling stages for the final impact analyses.
doi:10.1371/journal.pmed.1001250.g004

The Way Forward

In this exciting new era of HIV prevention technologies, C-RCTs will be used to test the hypothesis that combination HIV prevention, including expanded access to ART, can substantially reduce HIV incidence. Of particular relevance for the three planned C-RCTs is the observation that it may be challenging to observe a substantial reduction in HIV incidence (>40% reduction) over the 2- to 3-y duration of a trial unless the interventions are scaled up rapidly and the key populations are reached quickly. Models that reflect realistic delays in implementation and scale-up, as well as delays in the development of direct and indirect effects calibrated to the specific trial settings, will be particularly useful. These models will provide estimates of the effect size that can be expected at the end of the trial, the programmatic and implementation targets required to generate this effect, and the projected long-term impact. Ideally, the effect size should be chosen to be of public health relevance and to reflect long-term goals [5].

Given the challenges in scaling up interventions rapidly and the importance of these current trials, interim modelling analysis can provide a very useful and innovative tool to project the final intervention impact and to adopt mid-course corrections to accelerate scale-up and minimise the chance of having inconclusive trial results. However, the adaptive features of this design require careful statistical considerations so not to inflate the false

Key Points

- Cluster randomized controlled trials (C-RCTs) are currently planned to evaluate whether combination HIV prevention, including expanded access to ART, can substantially reduce HIV incidence at the population level in southern and eastern Africa.
- It may be challenging to observe a substantial reduction in HIV incidence (>40% reduction) over the 2- to 3-y duration of a C-RCT, unless the interventions are scaled up rapidly and the key populations are reached quickly.
- Mathematical models can and will be used to complement C-RCTs before, during, and after their completion to help plan, conduct, and interpret trial results and strengthen the evaluation of these interventions.
- Given the challenges in scaling up interventions rapidly and the importance of these current trials, we propose the innovative use of mathematical modelling to conduct interim analyses to modify or adapt an ongoing trial (in a carefully planned and prespecified manner) to reduce the likelihood of inconclusive trial outcomes, when interim HIV incidence data are not available.
- The interactive use of mathematical models during C-RCTs in a carefully preplanned fashion will also provide a unique opportunity to validate and refine model projections.

positive rate, which in turn requires modelling analysis to determine when that risk is outweighed by potential benefits.

The proposed modelling analyses will require collection of detailed data prior to and during the trial about the epidemiological context, and detailed information about the programmatic outcomes of each component will need to be available in a timely manner for key populations. Thus, it is critical that efficient data-capture systems are in place to allow linkage of HIV testing to the different services and the other components being modelled. There is also an emerging consensus that collecting detailed data characterising sexual networks will be important to interpret the results of the different trials effectively, especially if negative results are obtained. Efforts are currently ongoing to harmonise survey instruments across settings. The feasibility and added value of conducting complementary phylogenetic analyses to help understand transmission networks is also being considered.

Importantly, the interactive use of mathematical models during C-RCTs in a carefully preplanned fashion will not only be useful to demonstrate the use of models in designing, conducting, and interpreting C-RCTs, but will also provide a unique opportunity to validate and refine model projections. It will also test the usefulness of this modelling framework, which could then be used for C-RCTs designed to test prevention interventions for other infectious diseases with complex transmission dynamics such as malaria, tuberculosis, and neglected tropical infections.

Acknowledgments

This document draws on previous work and recent discussions on the conduct of C-RCTs of combination intervention held at the US National Institutes of Health/National Institute of Allergy and Infectious Diseases Workshop on Quantitative Methods to Advance the HIV Prevention Agenda, in Bethesda, Maryland (September 2011), and the HIV Modelling Consortium meeting "The Potential Impact of Treatment on HIV Incidence" in Stellenbosch, South Africa (November 2011) [11].

Author Contributions

Conceived and designed the experiments: MCB TBH. Performed the experiments: TBH MCB RA JFV. Analyzed the data: TBH MCB RA JFV. Wrote the first draft of the manuscript: TBH MCB BM. Contributed to the writing of the manuscript: MCB TBH BM JWE NSP RA JFV. ICMJE criteria for authorship read and met: MCB TBH BM JWE NSP RA JFV. Agree with manuscript results and conclusions: MCB TBH BM JWE NSP RA JFV. Produced the figures: RA TBH JWE. Produced tables: JFV MCB.

References

1. Joint United Nations Programme on HIV/AIDS (2011) World AIDS day report. How to get to zero: faster. smarter. better. Available: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216_WorldAIDSday_report_2011_en.pdf. Accessed 25 November 2011.
2. Alliance for Microbicide Development (2009) HIV/STI prevention research and development: October 2009 pipeline update of microbicide and PrEP candidates. Available: http://www.microbicide.org/uploads/3/1/2/1/3121935/microbicide_pipeline_update_1_oct_2009.pdf. Accessed 31 December 2011.
3. US President's Emergency Plan for AIDS Relief (2011 August) Guidance for the prevention of sexually transmitted HIV infections. Available: http://www.pepfar.gov/documents/organization/171303.pdf. Accessed 2 January 2012.
4. Schwartlander B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041.
5. Boily MC, Abu-Raddad L, Desai K, Masse B, Self S, et al. (2008) Measuring the public-health impact of candidate HIV vaccines as part of the licensing process. Lancet Infect Dis 8: 200–207.
6. Boily MC, Desai K, Masse B, Gumel A (2008) Incremental role of male circumcision on a generalised HIV epidemic through its protective effect against other sexually transmitted infections: from efficacy to effectiveness to population-level impact. Sex Transm Infect 84 (Suppl 2): ii28–ii34.
7. Brisson M, Edmunds WJ (2003) Economic evaluation of vaccination programs: the impact of herd-immunity. Med Decis Making 23: 76–82.
8. Donner A, Klar N (2000) Design and analysis of cluster randomization trials in health research. New York: Oxford University Press.
9. US Department of State (2011 September 14) PEPFAR announces largest study of combination HIV prevention. Available: http://www.state.gov/r/pa/prs/ps/2011/09/172389.htm. Accessed 2 January 2012.
10. Granich R, Gupta S, Suthar AB, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469.
11. HIV Modelling Consortium (2011) The potential impact of treatment on HIV incidence. Available: https://docs.google.com/viewer?a=v&pid=explorer&chrome=true&srcid=0B3AFjYLL-il3MGNhYjVkN2UtZmIwMy00ZWE4LWI4MDEtNTQ4NDVmNDNmYjRm&hl=en_US. Accessed 2 January 2012.
12. Padian NS, McCoy SI, Balkus JE, Wasserheit JN (2010) Weighing the gold in the gold standard: challenges in HIV prevention research. AIDS 24: 621–635.
13. Boily MC, Lowndes CM, Vickerman P, Kumaranayake L, Blanchard J, et al. (2007) Evaluating large-scale HIV prevention interventions: study design for an integrated mathematical modelling approach. Sex Transm Infect 83: 582–589.
14. Korenromp EL, Van Vliet C, Grosskurth H, Gavyole A, Van der Ploeg CP, et al. (2000) Model-based evaluation of single-round mass treatment of sexually transmitted diseases for HIV control in a rural African population. AIDS 14: 573–593.
15. Boily MC, Lowndes CM, Alary M (2000) Complementary hypothesis concerning the community sexually transmitted disease mass treatment puzzle in Rakai, Uganda. AIDS 14: 2583–2592.
16. Hallett TB, Garnett GP, Mupamberiyi Z, Gregson S (2008) Measuring effectiveness in community randomized trials of HIV prevention. Int J Epidemiol 37: 77–87.
17. Halloran ME, Struchiner CJ (1991) Study designs for dependent happenings. Epidemiology 2: 331–338.
18. Halloran ME (1998) Statistical issues in phase III human immunodeficiency virus vaccine trial design. J R Stat Soc Ser A Stat Soc 161: 265–272.
19. Alary M, Lowndes CM, Boily MC (2003) Community randomized trials for HIV prevention: the past, a lesson for the future? AIDS 17: 2661–2663.
20. Boily MC, Masse BR, Desai K, Alary M, Anderson RM (1999) Some important issues in the planning of phase III HIV vaccine efficacy trials. Vaccine 17: 989–1004.
21. Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N (2011) Mathematical models in the evaluation of health programmes. Lancet 378: 515–525.
22. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of

the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
23. Baggaley RF, Garnett GP, Ferguson NM (2006) Modelling the impact of antiretroviral use in resource-poor settings. PLoS Med 3: e124. doi:10.1371/journal.pmed.0030124
24. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
25. Johnstone-Robertson SP, Hargrove J, Williams BG (2011) Antiretroviral therapy initiated soon after HIV diagnosis as standard care: potential to save lives? HIV AIDS (Auckl) 3: 9–17.
26. Alsallaq R, Baeten J, Hughes J, Abu-Raddad L, Celum C, et al. (2012) Modeling the potential impact of a combination prevention based on community-wide HIV testing: KwaZulu Natal, South Africa [abstract]. 19th Conference on Retroviruses and Opportunistic Infections; 5–8 March 2012; Seattle, Washington, US. Available: http://www.retroconference.org/2012b/Abstracts/43187.htm. Accessed 31 May 2012.
27. Mehta SD, Gray RH, Auvert B, Moses S, Kigozi G, et al. (2009) Does sex in the early period after circumcision increase HIV-seroconversion risk? Pooled analysis of adult male circumcision clinical trials. AIDS 23: 1557–1564.
28. Cohen J (2011) HIV prevention. Halting HIV/AIDS epidemics. Science 334: 1338–1340.
29. Boily MC, Pickles M, Vickerman P, Buzdugan R, Isac S, et al. (2008) Using mathematical modelling to investigate the plausibility of attributing observed antenatal clinic declines to a female sex worker intervention in Karnataka state, India. AIDS 22 (Suppl 5): S149–S164.
30. Vickerman P, Watts C, Delany S, Alary M, Rees H, et al. (2006) The importance of context: model projections on how microbicide impact could be affected by the underlying epidemiologic and behavioral situation in 2 African settings. Sex Transm Dis 33: 397–405.
31. Dodd PJ, Garnett GP, Hallett TB (2010) Examining the promise of HIV elimination by 'test and treat' in hyperendemic settings. AIDS 24: 729–735.
32. Boily MC, Lowndes C, Alary M (2002) The impact of HIV epidemic phases on the effectiveness of core group interventions: insights from mathematical models. Sex Transm Infect 78 (Suppl 1): i78–i90.
33. Hallett TB, Singh K, Smith JA, White RG, Abu-Raddad LJ, et al. (2008) Understanding the impact of male circumcision interventions on the spread of HIV in southern Africa. PLoS ONE 3: e2212. doi:10.1371/journal.pone.0002212
34. Grassly NC, Garnett GP, Schwartlander B, Gregson S, Anderson RM (2001) The effectiveness of HIV prevention and the epidemiological context. Bull World Health Organ 79: 1121–1132.
35. Garnett GP, Anderson RM (1993) Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. Philos Trans R Soc Lond B Biol Sci 342: 137–159.
36. Robinson NJ, Mulder D, Auvert B, Whitworth J, Hayes R (1999) Type of partnership and heterosexual spread of HIV infection in rural Uganda: results from simulation modelling. Int J STD AIDS 10: 718–725.
37. Eaton JW, Hallett TB, Garnett GP (2011) Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS Behav 15: 687–692.
38. Boily MC, Alary M, Baggaley RF (2012) Neglected issues and hypotheses regarding the impact of sexual concurrency on HIV and sexually transmitted infections. AIDS Behav 16: 304–311.
39. Chen MI, Ghani AC, Edmunds J (2008) Mind the gap: the role of time between sex with two consecutive partners on the transmission dynamics of gonorrhea. Sex Transm Dis 35: 435–444.
40. Mishra S, Vickerman P, Pickles M, Ramesh BM, Washington R, et al. (2011) HIV prevention based on the static modes of transmission synthesis for two Indian districts: insights from dynamical modeling [abstract]. 19th Biennial Conference of the International Society for Sexually Transmitted Diseases Research; 10–13 July 2011; Quebec City, Canada. doi:10.1136/sextrans-2011-050108.584
41. Hayes RJ, Moulton LH (2009) Cluster randomized trials. Chapman & Hall/CRC Interdisciplinary Statistics Series. Boca Raton (Florida): CRC Press. 315 p.
42. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ (2008) Episodic sexual transmission of HIV revealed by molecular phylodynamics. PLoS Med 5: e50. doi:10.1371/journal.pmed.0050050
43. Pickles M, Foss AM, Vickerman P, Deering K, Verma S, et al. (2010) Interim modelling analysis to validate reported increases in condom use and assess HIV infections averted among female sex workers and clients in southern India following a targeted HIV prevention programme. Sex Transm Infect 86 (Suppl 1): i33–i43.
44. Hallett TB, Gregson S, Mugurungi O, Gonese E, Garnett GP (2009) Assessing evidence for behaviour change affecting the course of HIV epidemics: a new mathematical modelling approach and application to data from Zimbabwe. Epidemics 1: 108–117.
45. Blower S, Dowlatabadi H (1994) Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int Stat Rev 62: 229–243.
46. Johnson LF, Alkema L, Dorrington RE (2010) A Bayesian approach to uncertainty analysis of sexually transmitted infection models. Sex Transm Infect 86: 169–174.
47. Desai K, Boily MC, Masse B, Anderson RM (2006) Using transmission dynamics models to validate vaccine efficacy measures prior to conducting HIV vaccine efficacy trials. In: Abello J, Cormode G, editors. Discrete methods in epidemiology. Providence (Rhode Island): American Mathematical Society. pp. 139–163.
48. Emerson SS, Fleming TR (2010) Adaptive methods: telling "the rest of the story". J Biopharm Stat 20: 1150–1165.
49. Hsu HS, Gupta G (2010) An overview of the FDA draft guidance on adaptive design clinical trials [presentation]. MPCC Seminar Series. Available: http://www.fda.gov/downloads/BiologicsBloodVaccines/…/UCM209179.pdf. Accessed 16 March 2012.
50. Chow SC, Chang M (2008) Adaptive design methods in clinical trials—a review. Orphanet J Rare Dis 3: 11.
51. Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, et al. (2006) Adaptive designs in clinical drug development—an executive summary of the PhRMA Working Group. J Biopharm Stat 16: 275–283.
52. Zou GY, Donner A, Klar N (2005) Group sequential methods for cluster randomization trials with binary outcomes. Clin Trials 2: 479–487.
53. Cook RJ, Lawless JF (1996) Interim monitoring of longitudinal comparative studies with recurrent event responses. Biometrics 52: 1311–1323.
54. Meyer-Rath G, Over M (2012) HIV treatment as prevention: modelling the cost of antiretroviral treatment—state of the art and future directions. PLoS Med 9: e1001247. doi:10.1371/journal.pmed.1001247

OPEN ACCESS Freely available online

PLoS MEDICINE

Review

HIV Treatment as Prevention: Principles of Good HIV Epidemiology Modelling for Public Health Decision-Making in All Modes of Prevention and Evaluation

Wim Delva[1], David P. Wilson[2], Laith Abu-Raddad[3,4,5], Marelize Gorgens[6], David Wilson[6], Timothy B. Hallett[7], Alex Welte[1]*

1 South African Department of Science and Technology/National Research Foundation Centre for Excellence in Epidemiological Modelling and Analysis, University of Stellenbosch, Stellenbosch, South Africa, **2** Kirby Institute, University of New South Wales, Sydney, New South Wales, Australia, **3** Infectious Disease Epidemiology Group, Weill Cornell Medical College—Qatar, Cornell University, Qatar Foundation—Education City, Doha, Qatar, **4** Department of Public Health, Weill Cornell Medical College, Cornell University, New York, New York, United States of America, **5** Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America, **6** Global HIV/AIDS Program, The World Bank, Washington, District of Columbia, United States of America, **7** Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom

Abstract: Public health responses to HIV epidemics have long relied on epidemiological modelling analyses to help prospectively project and retrospectively estimate the impact, cost-effectiveness, affordability, and investment returns of interventions, and to help plan the design of evaluations. But translating model output into policy decisions and implementation on the ground is challenged by the differences in background and expectations of modellers and decision-makers. As part of the *PLoS Medicine* Collection "Investigating the Impact of Treatment on New HIV Infections"—which focuses on the contribution of modelling to current issues in HIV prevention—we present here principles of "best practice" for the construction, reporting, and interpretation of HIV epidemiological models for public health decision-making on all aspects of HIV. Aimed at both those who conduct modelling research and those who use modelling results, we hope that the principles described here will become a shared resource that facilitates constructive discussions about the policy implications that emerge from HIV epidemiology modelling results, and that promotes joint understanding between modellers and decision-makers about when modelling is useful as a tool in quantifying HIV epidemiological outcomes and improving prevention programming.

Introduction

In almost all areas of public health, mathematical models are used to provide quantification and insight that can inform decision-making. Epidemiological data can be collected about individuals, and clinical trials can measure individual-level effects in a selected study population (often under best-case circumstances), but public health decision-making requires an understanding of the dynamics of disease across a population under a variety of conditions. Mathematical modelling aims to unite knowledge and assumptions about behavioural dynamics, biology, costs, and constraints to generate estimates of impact and cost-effectiveness, and recommendations for resource allocation.

Models are especially useful in the case of infectious diseases, where they can estimate temporal changes in disease burden and treatment needs, and so underpin projections of the counterfactuals in some quasi-experimental impact evaluation designs, and power calculations for prospective experimental study designs. These are important applications, especially in contexts where empirical data are not available. Thus, models have increased in prominence over the last several years, including in establishing optimal responses to emerging pathogens [1] and influenza pandemics [2], examining the conditions for polio eradication [3] and malaria control [4], and making a case for restructuring investment in HIV programs [5,6].

Investigators from many different disciplines generate models, and the techniques and presentation formats employed have tended to follow a corresponding diverse set of conventions and presumptions. Meanwhile, those who rely on modelling output have highly varied needs and expectations from epidemiological modelling analyses. It is not uncommon for different models addressing very similar questions to produce—or appear to produce—widely different estimates [7], and thus a model's validity and ability to inform an important public health decision can be questioned.

Therefore, there is a need for constructive dialogue between "producers" and "consumers" of modelling results about a model's assumptions and structure, the policy implications of the results, and what further empirical and modelling studies should

Citation: Delva W, Wilson DP, Abu-Raddad L, Gorgens M, Wilson D, et al. (2012) HIV Treatment as Prevention: Principles of Good HIV Epidemiology Modelling for Public Health Decision-Making in All Modes of Prevention and Evaluation. PLoS Med 9(7): e1001239. doi:10.1371/journal.pmed.1001239

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Published July 10, 2012

Funding: The World Bank Global AIDS Program provided funding support to the writing committee. Two of the authors are employed by the World Bank, and the manuscript has been approved to be submitted with their World Bank affiliation. Apart from this approval, the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed here are those of the authors and not necessarily those of the funders, or numerous former and current collaborators, or members of the World Bank Prevention Science and Mathematical Modelling Reference Group, who provided feedback on the bare principles or earlier drafts of this work.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: ART, antiretroviral therapy

* E-mail: alexwelte@sun.ac.za

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

be planned. The World Bank Global HIV/AIDS Program, as a funder, coordinator, and evaluator of HIV prevention efforts, has become increasingly reliant on mathematical modelling and has initiated a modelling guidelines development process through its Prevention Science and Mathematical Modelling Reference Group, a panel of experts in HIV prevention, and modelling relating to HIV prevention, created and convened by the World Bank on the basis of individuals' publication records and institutional roles. In consultation with the reference group and other HIV modelling experts, we have developed a set of principles for the construction, reporting, and interpretation of HIV epidemiological models for public health decision-making on all aspects of HIV.

Development and Scope of the Recommendations

The nine principles, discussed below and summarised in Table 1, were initially identified during a number of discussions within the context of collaboration amongst the authors, within the HIV Modelling Consortium and the World Bank modelling guidelines production process. Written input on the nine principles was solicited from a wider group of modellers, including former and current collaborators. This was followed by a three-day work retreat of five of the authors, during which a first draft was produced, based on the authors' experience and other researchers' responses to the proposed core principles. The resulting draft was presented to a meeting of the World Bank Prevention Science and Mathematical Modelling Reference Group, and revised in light of feedback received.

Our focus complements more general reviews of modelling [8–10] and technical content in modelling textbooks [11,12]. The recommendations are intended for all HIV public health practitioners who rely on modelling research to make policy decisions, as well as those conducting the modelling research itself. They are not intended to be prescriptive, and hence should not be seen as a normative checklist against which to score the quality or validity of modelling studies. For instance, where mathematical models are used to construct a simple *conceptual* framework of behavioural, clinical, virological, and/or epidemiological dynamics, rather than to conduct research for public health decision-making, some of the recommendations in this article may not be applicable.

Principle 1: Clear Rationale, Scope, and Objectives

As in any scientific report, the rationale, scope, and objectives of a modelling study should be clearly stated. The reporting of a modelling study should include an explicit explanation for why epidemiological modelling, rather than another study design (e.g., systematic review, meta-analysis, quasi-experimental design, or a randomized controlled trial), is appropriate for the problem, the exact questions the work seeks to address, and the readership for which it is intended. This statement of rationale, scope, and objectives provides the criteria against which all modelling decisions should be judged, assists in framing the interpretation of the work, and should be referred to at key points throughout the write-up, to maintain the alignment of aims, model, results, and interpretation. Examples might be: "We aimed to generate estimates for the cost of rolling out a male circumcision programme in South Africa so that stakeholders can compare these costs against those of other possible interventions, and use the comparison to inform decisions about allocation of funding"; "We aimed to explore the extent to which HIV incidence rates can be influenced by changes in condom use among sex workers and their clients under different assumptions about sexual mixing patterns in concentrated HIV epidemics, so that recommendations can be made for data collection during the implementation of a condom distribution campaign".

For studies that aim to estimate the potential population-level impact of a given biomedical intervention, there are differences in emphasis in their purpose that should be clear from the outset and throughout the presented work. An important distinction is between investigation of the potential benefits of a hypothetical biomedical intervention that is currently in development but has unknown efficacy, and an intervention that has a proven efficacy, such as from a trial setting. Typically, the purpose of the first type of study is to estimate the population-level effectiveness of the hypothesized intervention and to identify key properties the intervention would need to have to be effective (such as for vaccines [13–15], microbicides [16,17], and chemoprophylaxis [18]), whereas the purpose of the second type of study is to guide targeted implementation of the intervention in real populations (such as deciding which populations should be circumcised first [19], or prioritised for treatment as prevention [20]). Another distinct form of modelling study is where an assessment is generated for the epidemiological impact of a previously implemented public health program [21].

Principle 2: Explicit Model Structure and Key Features

The model chosen for the analysis should be described completely and clearly (commonly in the form of an online technical appendix, ideally with the model's computer code made available), so that other investigators can reproduce its findings and projections. Justification for the choice of model (individual- versus population-based, stochastic versus deterministic, linear versus nonlinear) should be provided, along with a description of the model's structure and key features, with cross-references to the scope and objectives. A flow diagram, representing how individuals or subpopulations transition through the different demographic, behavioural, or clinical states in the model can be an excellent way to communicate the model's main structure.

The model structure, and the consequent key demographic, behavioural, biological, clinical, and epidemiological factors represented or omitted by the model, may affect the interpretation of the results. Certain biological or behavioural features of HIV transmission, prevention, and treatment may be at the core of the issue addressed by the model, and cannot be omitted. However, additional features that are irrelevant to the primary objectives of the analysis may obscure the main conclusions or may open unnecessary debate about the validity of parameter values that are not essential to interpretation of the model output [8]. Judging which features fall into which category may be informed by earlier research or explicit investigation, but is more commonly based on assumptions, which should at least be clearly stated. Furthermore, a mathematical model need not require an examination at all scales (e.g., within host, individual level, sexual network level, and population level); rather, scales to be included should be dictated by the objectives of the study (e.g., some models focus on within-host processes and thus must include the interaction between virus and immune cells, but models that focus on between-host transmission may not require detail at this scale). In general, the strength of the model should not be judged merely by the level of model detail and whether or not particular factors are included. Rather, the appropriateness of model detail and factors taken into

Table 1. Summary of principles of good HIV epidemiology modelling.

Principle	Model Producer Considerations	Model Consumer Considerations
Clear rationale, scope, and objectives	Are the rationale, scope, and objectives clearly stated?	Are the rationale, scope, and objectives understood?
	Is there a statement about why epidemiological modelling is appropriate for this problem?	Is epidemiological modelling appropriate for this problem?
Explicit model structure and key features	Is the model structure completely described, such that all analyses can be reproduced?	Is the model presented comprehensively, such that the inclusion/exclusion of any particular assumption or feature can be identified?
	Is there a description of key model features?	Is the justification for model structure/key assumptions reasonable, considering the primary rationale, scope, and objectives of the study?
	Has a justification for the model structure been provided?	
Well-defined and justified model parameters	Is there an understandable and complete listing of the model parameters, their values, and their justification?	Are the implicit inputs upon which the model predictions are made understood, and are they satisfactorily justified?
Alignment of model output with data	Are the model fitting, calibration, and validation approaches with respect to relevant data defined and justified?	Does the model produce, or fail to produce, outputs that can be compared to real world data, and does the model output reflect realistic conditions?
		Does the comparison with real world data increase confidence in the suitability of the model for the current enquiry?
Clear presentation of results, including uncertainty in estimates	Have the uncertainties been captured for all relevant factors included in the model?	Have the uncertainties been captured for all relevant factors included in the model?
	Is the key result of the study robust to that uncertainty?	Are the results sufficiently robust for confident decision-making, or is further analysis or data collection required?
	Are specific recommendations for new data analyses/collections appropriate?	
Exploration of model limitations	Are sufficient details provided about limitations of the study, specifically about model structure, parameterization, and application/generalisability?	Are the limitations of the model and its findings clearly understood, including the limits of applicability and generalisability?
		Considering the strength of the evidence, how are the model findings relevant for informing public health decision-making?
Contextualisation with other modelling studies	Have relevant previous studies been referenced and differences/similarities discussed?	Is there an understanding of the overarching conclusion(s) from modelling studies on the topic?
	Is it clearly specified whether a new result versus a confirmation/contradiction of a previous result is presented?	Are the general reasons (assumptions or underlying real world conditions) for why models differ in their conclusions understood?
Application of epidemiological modelling to health economic analyses	Where relevant, are understandable and appropriate estimates of epidemiological impact provided, such that health economic inferences can be made?	Can the model-based estimates be used to infer cost-effectiveness measures of relevant interventions or be extended to health economics?
		Is the degree of uncertainty in estimates relevant to cost-effectiveness understood, particularly with respect to the sensitivity of key parameters?
Clear language	Are model scenarios described in clear formal terms (separate from interpretations about reality) that facilitate technical understanding and evaluation?	Are there clear explanations of intended correspondences between inputs used in the model and key real world conditions such as epidemiological conditions, policy, and programmes?

doi:10.1371/journal.pmed.1001239.t001

account by the model should be assessed within the context of the scope and objectives.

Discussion of how the model structure could have influenced the results should always be included. Examples of formal evaluations of differently structured models addressing similar research questions but reaching different conclusions can be found in various branches within the infectious disease modelling field, e.g., in the modelling of chlamydia [22], influenza [23], and HIV epidemics [24]. It is often not feasible, in one article or within one modelling research group, to explore large differences in model structure, such as between deterministic population-based versus individual-based models. However, where possible, comparison between models is highly encouraged. For example, Johnson et al. [25] used two models in the same study to assess the impact of antiretroviral therapy (ART) and condom usage on HIV epidemics in South Africa, and Eaton et al. [7] discuss the implications of alternative model structures for estimating the potential impact of early initiation of ART on HIV incidence in hyperendemic settings. Such formal evaluations foster discussions of the reasons behind discrepancies in model predictions, and either pave the way for a consensus statement about the findings and conclusions that are most certain, or highlight key issues for further scientific enquiry.

Principle 3: Well-Defined and Justified Model Parameters

Another set of assumptions in a model concerns the values that are given to the parameters. Examples of parameters include the probability of HIV transmission per sex act for an individual on

ART, the fraction of patients still alive and on ART three years after ART initiation, and the annual population growth rate. It is essential for any modelling study to include a transparent listing of all model parameters, providing the following for each parameter: the name of the parameter; the mathematical symbol of the parameter (if appropriate); the meaning of the parameter in plain language; the value(s) assigned to the parameter (a point estimate and range/confidence interval as appropriate); and a contextual justification for used values, with references for the origins of the model parameter(s), and any relevant caveats (particularly important if more than one value for the model parameter exists or if the parameter is fit in the model or is derived from another modelling analysis).

This notion of justifying or formally "fitting" individual parameters—or a model in its entirety—to data covers many possibilities. As these also do not lie on a clear continuum from "rough heuristic/qualitative" to "formally rigorous and unbiased", some ad hoc critical evaluation is appropriate for the most important inputs into any model. All model fitting relies on the notion of the likelihood of observing a set of data. There are then various possible approaches to (1) maximising the likelihood, i.e., selecting the particular model within which the data are most consistent, or (2) performing a sensitivity analysis, i.e., identifying ranges of model parameters that are consistent with the data and determining the relative importance of each model parameter. Note that the "likelihood function" itself can capture multiple sources of randomness, such as the usually unavoidable incompleteness of sampling and random effects in population processes themselves.

Some parameters, such as the mother-to-child HIV transmission rate under a particular care regimen, can be more or less directly "measured" in an appropriate (typically randomized) study, using observation and standard robust biostatistical methods, but there may be subtle artefacts. For example, using logistic regression to identify the characteristics of individuals that are associated with an HIV infection or transmission event may be misleading in ways that are seldom systematically explored in routine application, beyond noting the potential for "residual confounding". A particular shape for a relationship between a predictor (such as viral load or age) and an outcome (transmission) is implicitly assumed, although it may be inappropriate—age in particular may correlate strongly with health status, but not necessarily monotonically.

For parameters where it is very difficult to obtain direct measurements, e.g., to capture behavioural dynamics such as risk reduction in the face of risk perception, heuristic parametrization may indicate which parameter sets are plausible and which are clearly at odds with data: a heuristically sensible model and a formally fitted model should be clearly distinguished, with sensitivity analyses where applicable.

Often the most important assumptions concern those specifying a simulated intervention, and it is recommended that these be prominently and exhaustively listed. For instance, if the intervention of interest relates to a policy change in ART, specifying a "coverage" and "efficacy" may not be enough: assumptions about enrolment rates, adherence, and retention, as well as behavioural characteristics (e.g., risk reduction or compensation) and demographic impacts (e.g., reduced mortality rates and increased size of the HIV-positive population) [7] may need to be made explicit. These specifications should be documented over the time period of the model simulation, and, where relevant, for different substrata of the modelled population. If the work is specific to a country, then it is helpful to involve relevant stakeholders in the decisions taken about parameter values, and this process should be described. Such documentation also assists when modelling findings are subsequently used to inform decision-making in that setting [26,27].

Principle 4: Alignment of Model Output with Data

Here the emphasis shifts to assessing the alignment of output from a particular epidemiological scenario model to data. Understanding the modelled scenarios produced, and relating these to data by back-fitting them to a model, naturally forms an important component of the evaluation and application of any model. It is particularly important to indicate whether, and to what extent, input parameters were chosen to maximise the correspondence of outputs to data, or whether correspondences emerged naturally from choosing externally justified inputs. Demonstrating that a model can reproduce observed patterns provides a certain level of reassurance that the model is capturing the system appropriately, and where models cannot demonstrate this, extreme caution should be taken in interpreting results.

The most desirable situation is when a model that has been fitted to some data (a training set) produces output in close correspondence with additional data (a testing set). There are two primary caveats to this approach: (1) fitting a smooth model to slowly varying data and extrapolating a little may be "too easy", and might indicate little about the suitability of the model, and (2) in key applications relevant to impact evaluation, asking the model to produce other independent data may be an unreasonable demand, tantamount to asking a model to predict future changes in the financial or political context. There may be deeper differences between the scenarios producing the training/testing datasets than can realistically be captured by a model—such as changes in treatment uptake or effects of improved treatment programmes on mortality.

While correspondence between models and data is reassuring and potentially useful—if not taken as absolute confirmation of the correctness of either model structure or parameter values—it is important to consider whether there are multiple ways to fit the data, and to realise that there may be scientific progress in a failure to fit data, either at all or without resorting to implausible values, ranges, or correlations of parameters. For example, simple (biological) models of ART cannot reproduce both the consistently strong reductions in patient viral loads and the inability to achieve viral eradication observed in the real world, without implausible "fine tuning" of individual subjects' treatment efficacy parameters into a narrow range. This situation diagnoses a model limitation, namely, the neglect of the fact that interactions between cells, drugs, and virions vary among compartments within the infected host.

The difficulties of "correctly" capturing a complex set of shifting context-defining processes impinge not only on the interpretation of correspondence between models and historical data, but also on the interpretation of the predictive component of scenarios. One useful application of modelling, when there are insufficient data to construct scenarios with conventional predictive credibility, is to pose questions such as what characteristic of a program would be required for certain goals to be achieved (e.g., what level of risk compensation, captured in a suitably clearly defined parameter, would be required to negate the risk reduction of a planned intervention).

Principle 5: Clear Presentation of Results, Including Uncertainty in Estimates

The output of any modelling study needs to be presented clearly, using explicitly defined metrics and with any deviance in the interpretation between the model metric and the real world analogue explained. The many assumptions involving the structure of the model, the parameter estimates, and the data will all have

uncertainties, and it is important to understand how these propagate to key model outputs. In some cases, uncertainty in a particular parameter will be benign—a result is reached irrespective of any credible assumption about that parameter—and this serves to increase confidence in the findings. In other cases, different credible values for a parameter (or model structure or interpretation of data) would lead to different conclusions, and this should be noted.

Uncertainties are best depicted as part of the modelling results presentation—either in tables or as part of the graphical output of the model. If sufficient information is available about inputs, computational techniques can manufacture a distribution for model outcomes, so that the main result can be given as a "credible interval". In addition to uncertainty analyses, formal sensitivity analyses of the importance of each model parameter in influencing the variability in model outcomes can be useful for identifying items for further data collection or investigation (see [28–30] for examples in HIV modelling). Bayesian melding approaches have also been used recently, and have the advantage that they integrate uncertainty analyses with model fitting: good examples in HIV transmission modelling include work by Alkema et al. [31] and Johnson et al. [32].

Principle 6: Exploration of Model Limitations

As Box and Draper [33] wrote, "Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful". It is necessary for modellers to provide a description of model limitations and for model consumers to appreciate the caveats and limitations of modelling studies when considering their results. Many limitations are due to the data that are available and used to parameterize modelling studies. Direct observation of some of the model parameters is often not feasible. This is especially true in the case of HIV, where transmission dynamics are dependent on sensitive and private aspects of human behaviour [34]. Modelling strategies address this challenge in part by fitting the model to data to yield estimates for the unknown parameters.

One thing that modellers may implicitly understand but that model consumers may not—and which therefore should always be made clear—is that capturing complex reality is not really the purpose of mathematical models. Practicality implies that one can never capture full dynamical structure, such as all conceivable population compartments, transition rules, or stochasticity. A mathematical model is a minimalist approach to representing the essential elements of reality that are necessary and sufficient for addressing a specific research question [35,36]. Models are often applied to specific settings, and so transferability of the predictions to other settings may be limited. Just as the findings of clinical trials can be subject to multiple interpretations, modelling studies similarly may have multiple interpretations, and even more readily admit various choices in emphasis, of which only a few receive a full airing in the investigators' report.

Some of the limitations of modelling studies can be addressed by uncertainty or sensitivity analyses as discussed above [28,37,38]. Probably the least appreciated mode by which limitations in models are addressed is by a comparative assessment of models and their predictions, similar to systematic reviews and meta-analyses of datasets. Recent examples of this kind of process include the male circumcision modelling consensus paper [19], a special edition of *Vaccine* that examined the potential impact of a partially effective vaccine [13], and model comparisons of the impact of ART on prevention presented in another article in the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections" [7].

Principle 7: Contextualisation with Other Modelling Studies

It is common for multiple modelling groups to attempt to address similar research questions but with different modelling approaches: using models that have been designed to describe different populations, involve different model structures, and make different parameter assumptions. Apparently conflicting results in the modelling literature may consequently lead to greater confusion for the consumers of models or to distrust in the use of models for decision-making. Therefore, it is necessary that interpretations of results are contextualised with previous modelling findings relevant to the topic. It should be made clear whether a new result is being presented or whether study findings concur with previously published results.

Meanwhile, journal editors should recognise the value of works that rigorously confirm or draw together previous findings. Review papers that summarise the modelling literature on a specific topic are highly useful (see the recent special issue on HIV epidemic modelling in *Current Opinion in HIV and AIDS* [39]). Also, papers that aim to present meta-analyses of model results (e.g., [24]) should be encouraged, as well as papers that compare modelling results to quasi-experimental results. Of even greater utility for policy-makers is the formulation of consensus documents that summarise conclusions from numerous modelling studies, and provide general conclusions in a single voice from the modelling community; this has been done for evaluations of circumcision interventions [19] and HIV vaccines [13], and this *PLoS Medicine* Collection on HIV treatment as prevention aims to move the field in that direction as well, although there is clearly much more to do [7,40].

Principle 8: Application of Epidemiological Modelling to Health Economic Analyses

A public health policy or programme decision-maker generally desires to take actions that will have maximal impact whilst minimising the amount of money required to achieve the health outcomes—based, for example, on estimates of either the maximum impact that can be achieved for a given amount of money, or the money needed to achieve specific set levels of impact. Therefore, the cost-effectiveness, affordability, and returns on investments of interventions are among the most important considerations in their potential implementation. HIV epidemic modelling studies often attempt to estimate the population-level impact associated with changes in programme or policy conditions, and hence estimate the denominator (effectiveness) in the incremental cost-effectiveness ratio. Ideally, such models should be designed to produce outputs amenable to recycling into analyses of cost implications and estimates of primary epidemiological effects that are understandable and relevant to decision-makers, such as the number of incident infections or deaths averted, quality-adjusted life years gained, or disability-adjusted life years averted. Effective assessment of affordability and cost-effectiveness may require different time horizons than those chosen in epidemiological modelling analyses, hence additional simulations may be necessary prior to attaching costs, benefits, and utilities to epidemiological model outputs.

There are numerous good examples of modelling studies that have provided outputs that are relevant for use in health economic calculations or that have been integrated into cost-effectiveness analyses [41–44]. Guidelines have been developed for the production, submission, and review of health economic analyses for *BMJ* [45]; some of the principles presented in those guidelines

align with those presented here. When modelling studies have the potential to be extended to health economic calculations, consideration of these health economic guidelines is encouraged.

Principle 9: Clear Language

A particular challenge that arises when using models to evaluate the impact of interventions is a lack of clarity around the intervention itself. Such a lack of clarity minimises the usefulness of results for policy-makers in deciding which interventions to prioritise. While modellers are usually keenly aware of the technical details of the model, the interpretation of model features—both in the input and output phase—is prone to oversimplification by both modellers and readers. It can be convenient but misleading to present a correspondence in the real world between an actual policy choice and future events. For instance, a write-up should highlight that what is modelled is a reduction in the proportion of "unprotected sex acts", which is not an intervention per se but could be the outcome of an intervention (e.g., an increase in condom distribution points or a targeted education campaign).

It is probably better to risk erring on the side of repetitiveness in efforts to keep focusing on precise model assumptions (qualitative and quantitative), and for consumers to process the model first on its own terms, before evaluating model scenarios in broad correspondence to reality and potential policy implications. At the same time, it is important that modellers use language that facilitates easy communication, without loss of precision and of key real world messages to consumers.

Conclusion

The issue of using models in decision-making is especially important for the field of HIV prevention, which has now reached a critical point. Just as spending on HIV has levelled off or declined [46], there have been several significant scientific breakthroughs, including the finding that ART can substantially reduce the infectiousness of infected individuals [47]. This finding immediately conjures a multitude of questions that can be best examined through mathematical modelling. Examples of specific questions within the field would include (1) whether programs should reallocate funding to treatment in response to the new data [48], (2) the probability of drug resistance emerging as a threat to the therapeutic effectiveness of treatment [49], and (3) how the impact of real programs can be scientifically measured [50].

Key Points

- Mathematical models are used to inform public health decision-making about many questions in the response to HIV epidemics, and here we present our recommendations for "best practices" for constructing, interpreting, and presenting such models.
- An overarching theme of our recommendations is that it is crucial for modellers to be explicit about the choices they make—about model structure, parameters, and model fitting and interpretation—and the reasoning behind their choices.
- Modellers need to make the limitations of their models clear, and model consumers (such as policy- and decision-makers) need to appreciate the caveats and limitations of modelling studies when considering their results.
- One of the least appreciated ways to address the limitations of models is through comparing the parameters, structure, and outputs of alternate models of the same processes.
- Especially useful are consensus documents that bring together conclusions from numerous modelling studies and summarise what researchers agree on and where uncertainty persists.

Further research questions are delineated in this *PLoS Medicine* Collection [40]. Our intention in compiling our recommendations is to help strengthen the support that mathematical models can provide in addressing such questions that are critical for setting research and intervention priorities for HIV.

Acknowledgments

The authors are grateful to Ms Britta Jewell for editorial assistance.

Author Contributions

Wrote the first draft of the manuscript: WD DPW LA-R TH AW. Contributed to the writing of the manuscript: WD DPW LA-R DW MG TH AW. ICMJE criteria for authorship read and met: WD DPW LA-R DW MG TH AW. Agree with manuscript results and conclusions: WD DPW LA-R DW MG TH AW. Contributed to the conceptualisation and drafting of the article: WD DPW LA-R DW MG TH AW.

References

1. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, et al. (2003) Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. Science 300: 1961–1966.
2. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, et al. (2009) Pandemic potential of a strain of influenza A (H1N1): early findings. Science 324: 1557–1561.
3. Grassly NC, Fraser C, Wenger J, Deshpande JM, Sutter RW, et al. (2006) New strategies for the elimination of polio from India. Science 314: 1150–1153.
4. Griffin JT, Hollingsworth TD, Okell LC, Churcher TS, White M, et al. (2010) Reducing Plasmodium falciparum malaria transmission in Africa: a model-based evaluation of intervention strategies. PLoS Med 7: e1000324. doi:10.1371/journal.pmed.1000324
5. Salomon JA, Hogan DR, Stover J, Stanecki KA, Walker N, et al. (2005) Integrating HIV prevention and treatment: from slogans to impact. PLoS Med 2: e16.doi:10.1371/journal.pmed.0020016
6. Schwartlander B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041.
7. Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med 9: e1001245. doi:10.1371/journal.pmed.1001245
8. Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N (2011) Mathematical models in the evaluation of health programmes. Lancet 378: 515–525.
9. Grassly NC, Fraser C (2008) Mathematical models of infectious disease transmission. Nat Rev Microbiol 6: 477–487.
10. Wilson DP, Garnett GP (2011) Introductions to recent developments in HIV epidemic modelling. Curr Opin HIV AIDS 6: 91–93.
11. Keeling MJ, Rohani P (2007) Modeling infectious diseases in humans and animals. Princeton: Princeton University Press.
12. Vynnycky E, White R (2010) An Introduction to Infectious Disease Modelling. Oxford: Oxford University Press.
13. Hankins CA, Glasser JW, Chen RT (2011) Modeling the impact of RV144-like vaccines on HIV transmission . Vaccine29: 6069–6071.
14. McLean AR, Blower SM (1993) Imperfect vaccines and herd immunity to HIV. Proc Biol Sci 253: 9–13.
15. Anderson R, Hanson M (2005) Potential public health impact of imperfect HIV type 1 vaccines. J Infect Dis 191 (Suppl 1): S85–S96.
16. Smith RJ, Bodine EN, Wilson DP, Blower SM (2005) Evaluating the potential impact of vaginal microbicides to reduce the risk of acquiring HIV in female sex workers. AIDS 19: 413–421.
17. Foss AM, Vickerman PT, Heise L, Watts CH (2003) Shifts in condom use following microbicide introduction: should we be concerned? AIDS 17: 1227–1237.
18. Desai K, Sansom SL, Ackers ML, Stewart SR, Hall HI, et al. (2008) Modeling the impact of HIV chemoprophylaxis strategies among men who have sex with men in the United States: HIV infections prevented and cost-effectiveness. AIDS 22: 1829–1839.

19. UNAIDS/WHO/SACEMA Expert Group on Modelling the Impact and Cost of Male Circumcision for HIV Prevention (2009) Male circumcision for HIV prevention in high HIV prevalence settings: what can mathematical modelling contribute to informed decision making? PLoS Med 6: e1000109. doi:10.1371/journal.pmed.1000109
20. Delva W, Eaton JW, Meng F, Fraser C, White RG, et al. (2012) HIV treatment as prevention: optimising the impact of expanded HIV treatment programmes. PLoS Med 9: e1001258. doi:10.1371/journal.pmed.1001258
21. Hallett TB, Garnett GP, Mupanberiyi Z, Gregson S (2008) Measuring effectiveness in community randomized trials of HIV prevention. Int J Epidemiol 37: 77–87.
22. Althaus CL, Turner KM, Schmid BV, Heijne JC, Kretzschmar M, et al. (2012) Transmission of Chlamydia trachomatis through sexual partnerships: a comparison between three individual-based models and empirical data. J R Soc Interface 9: 136–146.
23. Models of Infectious Diseases Agent Study (2011) MIDAS model profiles. Available: https://www.epimodels.org/midas/modelProfilesFull.do. Accessed 1 December 2011.
24. Baggaley RF, Ferguson NM, Garnett GP (2005) The epidemiological impact of antiretroviral use predicted by mathematical models: a review. Emerg Themes Epidemiol 2: 9.
25. Johnson LF, Hallett TB, Rehle TM, Dorrington RE (2012) The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. J R Soc Interface. E-pub ahead of print. doi:10.1098/rsif.2011.0826
26. Prestage GP, McCann P, Gray R, Hoare A, Down I, et al. (2011) The development of Australia's national syphilis action plan is based in interdisciplinary research findings. Int J Interdiscip Soc Sci 5: 239–261.
27. Gray RT, Heymer KJ, Hoare A, Kwon JA, Thein HH, et al. (2009) What impact might the economic crisis have on HIV epidemics in Southeast Asia? Curr HIV Res 7: 656–665.
28. Blower SM, Dowlatabadi H (1994) Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int Stat Rev 2: 229–243.
29. Wilson DP, Coplan PM, Wainberg MA, Blower SM (2008) The paradoxical effects of using antiretroviral-based microbicides to control HIV epidemics. Proc Natl Acad Sci U S A 105: 9835–9840.
30. Smith RJ, Okano JT, Kahn JS, Bodine EN, Blower S (2010) Evolutionary dynamics of complex networks of HIV drug-resistant strains: the case of San Francisco. Science 327: 697–701.
31. Alkema L, Raftery AE, Clark SJ (2007) Probabilistic projections of HIV prevalence using Bayesian melding. Ann Appl Stat 1: 229–248.
32. Johnson L, Dorrington R, Bradshaw D, Pillay-Van Wyk V, Rehle T (2009) Sexual behaviour patterns in South Africa and their association with the spread of HIV: insights from a mathematical model. Demogr Res 21: 289–340. doi:10.4054/DemRes.2009.21.11
33. Box G, Draper NR (1987) Empirical model-building and response surfaces. Oxford: John Wiley & Sons.
34. Abu-Raddad LJ, Schiffer JT, Ashley R, Mumtaz G, Alsallaq RA, et al. (2010) HSV-2 serology can be predictive of HIV epidemic potential and hidden sexual risk behavior in the Middle East and North Africa. Epidemics 2: 173–182.
35. Anderson RM, May RM (1991) Infectious disease of humans: dynamics and control. Oxford: Oxford University Press.
36. Abu-Raddad LJ, Patnaik P, Kublin JG (2006) Dual infection with HIV and malaria fuels the spread of both diseases in sub-Saharan Africa. Science 314: 1603–1606.
37. Marino S, Hogue IB, Ray CJ, Kirschner DE (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol 254: 178–196.
38. Abu-Raddad L, Longini IM (2008) No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. AIDS 22: 1055–1061.
39. Wilson DP, Garnett GP (2011) Introduction to recent developments in HIV epidemic modeling. Curr Opin HIV AIDS 6: 91–93.
40. The HIV Modelling Consortium Treatment as Prevention Editorial Writing Group (2012) HIV treatment as prevention: models, data, and questions—Towards evidence-based decision-making. PLoS Med 9: e1001259. doi:10.1371/journal.pmed.1001259
41. Long EF, Brandeau ML, Owens DK (2010) The cost-effectiveness and population outcomes of expanded HIV screening and antiretroviral treatment in the United States. Ann Intern Med 153: 778–789.
42. Anderson JS, Wilson D, Templeton DJ, Grulich A, Carter R, et al. (2009) Cost-effectiveness of adult circumcision for HIV prevention in men who have sex with men in a resource-rich setting. J Infect Dis 200: 1803–1812.
43. Vickerman P, Kumaranayake L, Balakireva O, Guinness L, Artyukh O, et al. (2006) The cost-effectiveness of expanding harm reduction activities for injecting drug users in Odessa, Ukraine. Sex Transm Dis 33 (Suppl 10): S89–S102.
44. Alistar SS, Owens DK, Brandeau ML (2011) Effectiveness and cost effectiveness of expanding harm reduction and antiretroviral therapy in a mixed HIV epidemic: a modeling analysis for Ukraine. PLoS Med 8: e1000423. doi:10.1371/journal.pmed.1000423
45. Drummond MF, Jefferson TO (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ 313: 275–283.
46. Kaiser Family Foundation (2011) International AIDS assistance from donor governments: commitments & disbursements, 2002–2010. Available: http://facts.kff.org/chart.aspx?ch=946. Accessed 25 November 2011.
47. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505.
48. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57.
49. Wagner BG, Kahn JS, Blower S (2010) Should we try to eliminate HIV epidemics by using a 'test and treat' strategy? AIDS 24: 775–776.
50. Hallett TB, White PJ, Garnett GP (2007) Appropriate evaluation of HIV prevention interventions: from experiment to full-scale implementation. Sex Transm Infect 83 (Suppl 1): i55–i60.

HIV Treatment as Prevention: Systematic Comparison of Mathematical Models of the Potential Impact of Antiretroviral Therapy on HIV Incidence in South Africa

Jeffrey W. Eaton[1]*, Leigh F. Johnson[2], Joshua A. Salomon[3], Till Bärnighausen[3,4], Eran Bendavid[5], Anna Bershteyn[6], David E. Bloom[3], Valentina Cambiano[7], Christophe Fraser[8], Jan A. C. Hontelez[4,9,10], Salal Humair[3,11], Daniel J. Klein[6], Elisa F. Long[12], Andrew N. Phillips[7], Carel Pretorius[13], John Stover[13], Edward A. Wenger[6], Brian G. Williams[14], Timothy B. Hallett[1]

1 Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **2** Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa, **3** Harvard School of Public Health, Boston, Massachusetts, United States of America, **4** Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Mtubatuba, South Africa, **5** Department of Medicine, Stanford University, Stanford, California, United States of America, **6** Intellectual Ventures Laboratory, Bellevue, Washington, United States of America, **7** Research Department of Infection & Population Health, University College London, London, United Kingdom, **8** Medical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **9** Erasmus University, Rotterdam, Netherlands, **10** Department of Primary and Community Care, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands, **11** School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan, **12** Yale University, New Haven, Connecticut, United States of America, **13** Futures Institute, Glastonbury, Connecticut, United States of America, **14** South African Centre for Epidemiological Modelling and Analysis, Stellenbosch, South Africa

Abstract

Background: Many mathematical models have investigated the impact of expanding access to antiretroviral therapy (ART) on new HIV infections. Comparing results and conclusions across models is challenging because models have addressed slightly different questions and have reported different outcome metrics. This study compares the predictions of several mathematical models simulating the same ART intervention programmes to determine the extent to which models agree about the epidemiological impact of expanded ART.

Methods and Findings: Twelve independent mathematical models evaluated a set of standardised ART intervention scenarios in South Africa and reported a common set of outputs. Intervention scenarios systematically varied the CD4 count threshold for treatment eligibility, access to treatment, and programme retention. For a scenario in which 80% of HIV-infected individuals start treatment on average 1 y after their CD4 count drops below 350 cells/μl and 85% remain on treatment after 3 y, the models projected that HIV incidence would be 35% to 54% lower 8 y after the introduction of ART, compared to a counterfactual scenario in which there is no ART. More variation existed in the estimated long-term (38 y) reductions in incidence. The impact of optimistic interventions including immediate ART initiation varied widely across models, maintaining substantial uncertainty about the theoretical prospect for elimination of HIV from the population using ART alone over the next four decades. The number of person-years of ART per infection averted over 8 y ranged between 5.8 and 18.7. Considering the actual scale-up of ART in South Africa, seven models estimated that current HIV incidence is 17% to 32% lower than it would have been in the absence of ART. Differences between model assumptions about CD4 decline and HIV transmissibility over the course of infection explained only a modest amount of the variation in model results.

Conclusions: Mathematical models evaluating the impact of ART vary substantially in structure, complexity, and parameter choices, but all suggest that ART, at high levels of access and with high adherence, has the potential to substantially reduce new HIV infections. There was broad agreement regarding the short-term epidemiologic impact of ambitious treatment scale-up, but more variation in longer term projections and in the efficiency with which treatment can reduce new infections. Differences between model predictions could not be explained by differences in model structure or parameterization that were hypothesized to affect intervention impact.

Please see later in the article for the Editors' Summary.

Citation: Eaton JW, Johnson LF, Salomon JA, Bärnighausen T, Bendavid E, et al. (2012) HIV Treatment as Prevention: Systematic Comparison of Mathematical Models of the Potential Impact of Antiretroviral Therapy on HIV Incidence in South Africa. PLoS Med 9(7): e1001245. doi:10.1371/journal.pmed.1001245

Academic Editor: John Bartlett, Duke University Medical Center, United States of America

Received January 2, 2012; **Accepted** May 10, 2012; **Published** July 10, 2012

Funding: The HIV Modelling Consortium is supported by a grant from the Bill and Melinda Gates Foundation to Imperial College London. JWE thanks the Marshall Aid and Commemoration Commission for scholarship support. EFL acknowledges support from the U.S. National Institute on Drug Abuse (R01-DA15612). Till Bärnighausen was supported through grant R01 HD058482-01 from the National Institute of Child Health and Development, US National Institutes of Health (NIH); 1R01MH083539-01 from the National Institute of Mental Health, NIH; World Bank; and the Wellcome Trust; and he received funding for research on economic evaluation of Dengue vaccination from Sanofi-Pasteur and for research on economic evaluation of human papilloma virus vaccination from Merck. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: I have read the journal's policy and have the following conflicts: JAS is a member of the *PLoS Medicine* Editorial Board. TB, DEB, and SH received a grant from the World Bank Global HIV/AIDS program through the Economics Reference Group for the initial development of the BBH model.

Abbreviations: ART, antiretroviral therapy

* E-mail: jeffrey.eaton@imperial.ac.uk

Provenance: Submitted as part of a sponsored Collection; externally reviewed.

Introduction

There has recently been increasing interest in expanding provision of antiretroviral therapy (ART) as a tool for reducing the spread of HIV in generalised epidemics in sub-Saharan Africa [1–5]. As momentum gathers for "HIV treatment as prevention", there is an urgent need to understand how ART might contribute to averting HIV transmissions, in addition to its direct benefits in reducing morbidity and mortality amongst treated patients. Mathematical modelling has supplied critical insights to discussions about treatment as prevention by providing a framework for combining information about the relationship between an infected individual's viral load and HIV transmissibility [6,7], the reduction in a host's HIV viral load when on ART [8,9], and the population-level contact structure over which HIV is transmitted [10,11].

The idea of using medicines that suppress viral concentrations to reduce transmission of infection was posed almost as soon as the first HIV drugs were developed [12,13]. Early models of the impact of ART focused on the opposing effects of reduced transmissibility and extended survival on new HIV infections, and whether associated increases in sexual risk behaviour would negate the prevention benefits of ART [10,12,14–23]. Since then, longitudinal observational data and one randomized controlled trial have demonstrated substantial reductions in the risk of heterosexual HIV transmission when the infective partner is virally suppressed [24–28], and continued follow-up of individuals receiving ART has confirmed the durability of viral suppression [29], including in sub-Saharan Africa [30,31]. At the same time, there have been tremendous improvements in access to treatment in sub-Saharan Africa [32]. More recent modelling has shown more optimism about the potential for treatment to reduce new HIV infections in this region, with much work focused on the setting of South Africa, home to one in six people living with HIV globally [33].

Perhaps the most provocative of these modelling efforts has been the study by Granich and colleagues suggesting that a strategy involving annual testing and immediate treatment for all HIV-infected individuals, combined with other interventions, could eliminate HIV by the year 2050 [34]. Wagner and Blower implemented the same model but used different assumptions about treatment uptake amongst asymptomatic infected individuals that they characterised as being more realistic, and concluded that elimination would not be possible [35]. Kretzschmar et al. highlighted how choices in model structure affect the epidemic dynamics and intervention impact [36]. Dodd et al. showed that the potential for treatment to eliminate HIV depends on the patterns of sexual mixing in the population [11]. An age-structured model by Bacaër et al. found that elimination might be possible with less frequent testing than proposed by Granich et al., given recent epidemic trends and increases in condom usage [37]. Bendavid et al. used a microsimulation model to highlight that, in addition to increasing HIV testing, improving linkage to and retention in care are essential to achieving maximal benefits of test-and-treat interventions [38].

Other models have focused on the potential prevention benefits of providing treatment in line with current therapeutic guidelines. Eaton et al. estimated that 60 to 90 new infections could be averted for every 1,000 additional persons treated with CD4 cell count below 350 cells/μl (the current World Health Organization recommendation for when to start treatment [39]), depending on how well patients on treatment are retained in care [40]. The Goals model, used in the evaluation of the new UNAIDS Investment Framework, found that a US$46.5 billion incremental investment over the years 2011 to 2020, incorporating expanding access to ART, could avert 12.2 million new infections and 7.4 million deaths globally over that period [41]. Using a microsimulation model of the HIV epidemic in KwaZulu-Natal Province, Hontelez et al. found that expanding access to ART from those with CD4 cell count ≤200 cells/μl to those with ≤350 cells/μl required 28% more patients to receive treatment, but amounted to only a 7% increase in annual investment [42]. Cumulative net costs broke even after 16 y.

Models have also sought to understand the impact of past and current treatment policies; Johnson et al. used the ASSA2003 and STI-HIV Interaction models to assess the relative contributions of increased condom usage and ART scale-up to the declines in HIV incidence in South Africa up to 2008 [43]. Finally, other mathematical models have been used for short-term projections as a basis for power calculations for community-randomized trials of treatment as prevention [44].

Each of these models has predicted dramatic epidemiologic benefits of expanding access to ART, but models appear to diverge in their estimates of the possibility of eventually eliminating HIV using ART, the cost-effectiveness of increasing the CD4 threshold for treatment eligibility, and the benefits of immediate treatment compared to treatment based on the current World Health Organization eligibility guidelines. Directly comparing the models' predictions is challenging because each model has been applied to a slightly different setting, has used different assumptions regarding other interventions, has been used to answer different questions, and has reported different outcome metrics.

In this study we seek to understand the extent to which diverse mathematical models agree on the epidemiological impact of expanded access to ART by simulating the same set of intervention scenarios across the models and focusing on standardised outputs. The intervention scenarios are designed to be simple enough to be consistently implemented across different types of models in order to control several aspects of the treatment programme and isolate the effects of model structure, parameters, and assumptions about the underlying epidemic on estimates of intervention impact. The purpose of this study is not to make predictions about the impact of any particular intervention in any specific setting, but rather to better characterise the array of mathematical models being used to inform policy about treatment as prevention in hyperendemic settings such as South Africa.

Methods

Study Design

Literature and reports of meetings on related topics were reviewed in August 2011, and researchers who had previously developed mathematical models of the potential epidemiological impact of expanded access to ART, calibrated to the South African epidemic setting, were invited to participate in the model comparison exercise by simulating a standardised set of ART scale-up scenarios. Three aspects of the treatment programme were systematically varied: the CD4 threshold for treatment eligibility, access to treatment for those eligible, and the retention of patients on treatment. The timing of ART introduction and the rate at which individuals start treatment after becoming eligible were also standardised. The impact of an intervention was measured by comparing the number of new infections in the intervention scenario with that in a counterfactual epidemic simulation in which no ART is provided within the same model population. The counterfactual of no ART was chosen so that comparison between models would be independent of assumptions

about the historic growth in ART uptake. As such, the results should not be interpreted as estimates of the future impact of treatment compared to current patterns of ART coverage, but can be generally taken as estimates of the overall net impact of treatment in a hypothetical scenario that assumes rapid ART scale-up and a homogenous rate of ART initiation across all ART-eligible adults. Although different models may incidentally have been calibrated using the same data, no standardisation was imposed on the specific epidemiologic data used for model calibration or on the calibration procedure itself in this exercise.

Mathematical Models

Twelve groups accepted the invitation to participate in the model comparison exercise. The collection of models encompasses a wide range of model structures, mechanisms for representing HIV transmission and disease progression, overall levels of complexity, and detail in the characterisation of treatment programmes. Table 1 summarises the names, authors, and key references for each model, and compares aspects of model structure. Four of the models are agent-based microsimulation models (i.e., models that track the behaviour and infection status of individual people) and use random-number generators to simulate particular events such as a new partnership formation or transmission events. The remaining eight models are deterministic compartmental models that stratify the population into groups according to each individual's characteristics and HIV infection status and use differential or difference equations to track the rate of movement of individuals between these groups. One of the models, the BBH model, solves the differential equations analytically, while the others numerically evaluate the differential equations. Ten of the models explicitly simulate both sexes and heterosexual HIV transmission, and six of the models include some form of age structure, although the extent to which age affects the natural history of HIV, the risk of HIV acquisition, and the risk of HIV transmission varies amongst these. All of the models simulate the national HIV epidemic in South Africa except for the STDSIM model, which simulates the higher prevalence Hlabisa subdistrict of KwaZulu-Natal Province, South Africa. Box 1 gives further comparative description of the structures and parameterization of the mathematical models.

Intervention Scenarios

Three different CD4 cell count thresholds for treatment eligibility were considered: CD4 count $\leq$200 cells/μl, CD4 count $\leq$350 cells/μl, and all HIV-infected individuals. In each eligibility scenario, treatment initiation was simulated under the assumption that all eligible individuals had equal access, without prioritisation for any subpopulations. It was further assumed that eligible individuals with access to the intervention would initiate ART at a constant rate after reaching eligibility, such that average time from eligibility to treatment initiation would be 1 y.

Treatment access was defined as the proportion of eligible individuals who eventually initiate treatment. For example, 60% access and eligibility at CD4$\leq$350 cells/μl implies that 60% of individuals will initiate treatment, on average 1 y after their CD4 count drops below 350 cells/μl, while 40% will never access treatment. Seven levels of treatment access were evaluated: 50%, 60%, 70%, 80%, 90%, 95%, and 100%.

Retention was defined as the percentage of individuals remaining on treatment after 3 y, excluding from both the numerator and the denominator those who had died while on treatment. The levels of retention were 75%, 85%, 95%, and 100% (no dropout), with individuals dropping out from treatment at a constant rate such that the desired level of retention was achieved at the 3-y time point. The prognosis and future treatment options for individuals who dropped out from treatment were not standardised.

Intervention Scale-Up

ART was assumed to be introduced into the population from the beginning of year 2012, with no treatment provision prior to this (in contrast to the rapid scale-up of treatment that has actually occurred prior to 2012 in South Africa). Intervention scale-up was immediate—a fraction (corresponding to the specified level of ART access) of individuals already eligible for treatment at the start of the intervention period were assumed to initiate treatment at a constant rate from that point, along with individuals who became eligible for treatment after the start of the intervention period.

Output Metrics

The measures of intervention impact were the percentage reduction in HIV incidence rate among adults (aged 15–49 y) in the ART scenario versus the no-ART counterfactual, the cumulative number of person-years of ART provided since the introduction of ART, and the cumulative number of person-years of ART provided per infection averted as a measure of the "efficiency" with which ART prevents infections. The percentage reduction in incidence was defined by calculating the difference in the adult HIV incidence rate between the intervention and no-ART counterfactual in the same year and dividing this by the incidence rate in the counterfactual scenario. The number of person-years of ART provided per infection averted was calculated by dividing the cumulative number of person-years of ART by the difference between the number of new adult infections since year 2012 in the intervention and the counterfactual scenario. Each of these metrics was reported at the midpoints of the years 2020 and 2050. Most of the models included in this study were not designed with the intention of making realistic projections to year 2050, but these results were included to gain some insight into the long-term dynamics of the models.

In addition to these measures of intervention impact, each model reported the HIV prevalence and HIV incidence rate amongst males and females aged 15–49 y for the no-treatment counterfactual simulation and the total size of the adult population (age 15 y and older). Each model also produced the proportion of the HIV-infected population in each CD4 count category ($\leq$200, 200–350, and $>$350 cells/μl) and in early HIV infection in year 2012, and the proportion of HIV transmissions from individuals in each category.

The Eaton and STI-HIV Interaction models report posterior means and 95% credible intervals for model outcomes of interest (see Box 1). The Bendavid model completed simulations only for 50%, 80%, and 100% access, and 75%, 85%, and 100% retention scenarios, and only simulated to year 2040, so results for this model are reported for the year 2040 where other model results are reported for year 2050. The BBH model completed simulations only for the 85% and 100% retention scenarios. The Granich model did not simulate ART for the CD4$\leq$200 cells/μl eligibility threshold, while the STI-HIV Interaction model did not simulate ART eligibility for all HIV-infected individuals. As a result of these models not completing all intervention scenarios and outputs, some analyses include only a subset of the models. To maximise comparability, the 40% reduction in transmission due to combination with other preventive interventions assumed by Granich and colleagues in [34] was not incorporated here.

Table 1. Description of mathematical models.

Model Name [Key References]	Model Authors	Model Type	Model Calibration	Population Structure	Sexual Mixing	Variation in Infectiousness over Course of Infection	Increased Male-to-Female Transmission?	Reduced Transmission on ART	Behaviour Change in 2000s
BBH	Till Bärnighausen, David Bloom, Salal Humair	Deterministic (analytically derived)	Initialized to epidemic state in year 2009; HIV incidence generated using national data on HIV prevalence, mortality, ART, and sexual behaviour	Two-sex, age 15–49 y	Homogeneous heterosexual mixing	Early infection, latent infection, late-stage/AIDS	3 times greater	96.5%	Not applicable
Bendavid [38]	Eran Bendavid	Microsimulation	Calibrated to current epidemic state in year 2012; select parameters estimated by scanning ranges from literature.	Two-sex, age-structured	Short- and long-term partners; heterogeneity in number of partners; decrease with age; short partners assortative by 5-y age group	Early infection, then according to VL (partner VL not tracked, randomly sampled from population)	No, but includes circumcision	91% (mediated by VL)	Not applicable
CD4 HIV/ART (CD4 Model of HIV and ART) [77]	John Stover, Carel Pretorius	Deterministic	Incidence curve imported from Spectrum model [78] projection for South Africa 2009 national estimates	Single-sex, age 15+ y	Sexual mixing not explicitly modelled; infection by multiplying fixed force of infection by current HIV prevalence	Early infection (×9.2), asymptomatic, symptomatic (×7.3)	Not applicable	96%	Not applicable (incidence curve input from Spectrum)
Eaton	Jeffrey Eaton, Timothy Hallett, Geoffrey Garnett	Deterministic	Sex-specific age (15–49 y) HH prevalence (HSRC '02, '05, '08) and ANC prevalence '90–'08 estimated using Bayesian framework	Two-sex, age 15–49 y (sexually active) and age 50+ y (not sexually active)	Three sexual risk groups, with partially assortative mixing by risk group	Early infection (×38), CD4>350 (×0.61), 350>CD4>200, 200>CD4>100 (×3.75), CD4≤100 (×0.71)	No	92%	Reduction in unprotected sexual contacts over ~1999 to 2011 (timing and amount estimated)
EMOD	Daniel Klein, Anna Bershteyn, Edward Wenger, Karima Nigmatulina, Philip Eckhoff	Microsimulation	HIV prevalence time series (ANC '90–'09) and by age and sex (HSRC '08); sexual behaviour informed by Africa Centre for Health and Population Studies [79]	Two-sex, age-structured	Transitory, informal, and marital relationships; heterogeneity in propensity for each type of relationship by age and sex	Early infection (×26), asymptomatic, AIDS (×7.2)	No, but includes male circumcision	96%	Increase in condom usage; most change occurs between 1999 and 2009
Fraser	Christophe Fraser	Deterministic	UNAIDS age 15–49 y HIV prevalence estimates, fit by least squares	Two-sex, age 15–49 y	Three sexual risk groups, with partially assortative mixing by risk group	Early infection (×26, adjusted for partner duration), CD4>200, CD4≤200 (×2.4)	No, but 76% of males are circumcised (based on Western Cape)	90%	No
Goals [41,80]	Carel Pretorius, John Stover	Deterministic	Calibrated to match time series in HIV prevalence from Spectrum projection	Two-sex, age 15–49 y	Not sexually active, low, medium, high (CSW and clients) risk, plus IDU and MSM; mixing perfectly assortative by risk group except low risk mix with CSW client	Early infection (×28), asymptomatic, symptomatic (×7.3)	1.4 times greater	92%	Increase in condom usage over 1996 to 2009
Granich [34]	Brian Williams, Reuben Granich	Deterministic	Calibrated to annual national age 15–49 y prevalence estimates	Single-sex, age 15–49 y	Homogeneous	No	Not applicable	99%	No

Table 1. Cont.

Model Name [Key References]	Model Authors	Model Type	Model Calibration	Population Structure	Sexual Mixing	Variation in Infectiousness over Course of Infection	Increased Male-to-Female Transmission?	Reduced Transmission on ART	Behaviour Change in 2000s
HIV Portfolio [81]	Elisa Long, Margaret Brandeau, Douglas Owens	Deterministic	Initialized to current epidemic state in year 2011	Two-sex, age 15–49 y	Homogeneous	Early infection (~×5), CD4>350, 350>CD4 >200 (~×1.6), CD4 ≤200 (~×2)	~1.5 times greater, plus circumcision	90%	Not applicable
STDSIM [42,82,83]	Jan Hontelez, Sake de Vlas, Frank Tanser, Roel Bakker, Till Bärnighausen, Marie-Louise Newell, Rob Baltussen, Mark Lurie	Microsimulation	Calibrated to the Hlabisa subdistrict of KwaZulu-Natal using data from the Africa Centre for Health and Population Studies [79]	Two-sex, age-structured	Marriages, casual partnerships, and commercial sex; heterogeneity in propensity to form each type or relationship; changes with age	Early infection (×15), asymptomatic, symptomatic (×3), AIDS (×7.5)	2 times greater, plus circumcision	92%	Increase in condom usage and improvement in STI treatment over 1995 to 2003
STI-HIV Interaction [43,84]	Leigh Johnson	Deterministic	Age-specific household prevalence (HSRC '02, '05, '08) and age-specific ANC prevalence ('90–'08) estimated using Bayesian framework	Two-sex, age-structured (5-y age groups)	High and low risk assortatively mixing; further stratified by short- and long-term partnerships, plus CSW	Early infection (×10), asymptomatic, pre-AIDS (×2.5), AIDS (×5)	~2.5 times greater	90% (range: 78%–98%)	Increase in condom usage; most of increase occurs between 1995 and 2005
Synthesis Transmission [66]	Valentina Cambiano, Andrew Phillips, Deenan Pillay, Jens Lundgren, Geoff Garnett	Microsimulation	Calibrated using national HH survey data (HSRC '02, '05, '08)	Two-sex, age-structured	Primary partner and short-term partners; four different risk groups for propensity for short-term partners; semi-assortative by age	Early infection, then transmission determined by VL (primary partner's VL tracked, short-term sampled from population VL distribution)	~1.5 times greater	Determined by adherence and viral load	Reduction in number of unprotected partners over period 1996 to 2008

ANC, antenatal clinic; CSW, commercial sex worker; HH, household; HSRC, South African Human Sciences Research Council; IDU, injecting drug user; MSM, men who have sex with men; STI, sexually transmitted infection; VL, viral load.

doi:10.1371/journal.pmed.1001245.t001

Box 1. Comparative Description of Mathematical Models

This box elaborates on the comparison of aspects of the models' structure, assumptions, and parameterization presented in Table 1. Specific details about the structure and implementation of each of the models are available in the references included in Table 1 or from the HIV Modelling Consortium (http://www.hivmodelling.org/plos-medicine-special-collection).

Many of the models allow individuals to have different propensities for sexual risk behaviour. Each of the microsimulation models allows individuals to have both long-term (or marital) partnerships and short-term (or informal or casual) partnerships that are different in duration, and individuals have heterogeneous propensities to form short-term partnerships. In the STDSIM model a proportion of the population engages in commercial sex work partnerships; in the EMOD model a proportion can have transitory partnerships, a third partnership type that is shorter than a casual partnership. Among the microsimulation models, EMOD and STDSIM explicitly simulate the sexual partnership network, while the Bendavid and Synthesis Transmission models calculate the risk of acquiring HIV for an individual in a partnership by sampling the distribution of viral load across, respectively, the entire population and potential partners.

The deterministic models assume that sexual contacts occur instantaneously. The BBH, Granich, and HIV Portfolio models assume that all individuals form new contacts at the same rate and mix homogeneously. The other deterministic models stratify the population into risk groups that form new contacts at different rates (Eaton, Fraser, and Goals: three groups; STI-HIV Interaction: two groups). The STI-HIV Interaction and Goals models additionally include commercial sex workers, and the Goals model includes transmission among men who have sex with men and injecting drug users. The STI-HIV Interaction model separates both the low- and high-risk groups into those with short-term or long-term partnerships or both. The Eaton, Fraser, and STI-HIV Interaction models all include a degree of ''assortative'' mixing (preferential partnership formation with those in the same risk group), and all partnerships are formed in the same risk group in the Goals model, except for low-risk men and women who are married to high-risk partners. The CD4 HIV/ART model does not explicitly model sexual mixing but rather calculates the number of new HIV infections by multiplying the current number of HIV-infected adults by a fixed force of infection calculated from the Spectrum model projection for South Africa.

All of the models except for the Granich model simulate different stages of HIV infection that affect the transmissibility of an individual, including a period of elevated infectiousness during the first few weeks of infection and increased transmission during later stage infection. Parameters governing the relative transmissibility during early infection are based principally on two sources: a meta-analysis of HIV transmission per coital act by Boily et al. [68], which estimated a 10-fold increase in transmission relative to asymptomatic infection (BBH, CD4 HIV/ART, Goals, and STI-HIV Interaction), or a reanalysis of data from Rakai, Uganda [70], by Hollingsworth et al. [69], which estimated a 26-fold increase (Eaton, EMOD, Fraser, and Synthesis Transmission). Relative transmissibility after the early stage is according to clinical stage (asymptomatic and AIDS: BBH, CD4 HIV/ART, EMOD, Goals; asymptomatic, pre-AIDS, and AIDS: STDSIM, STI-HIV Interaction) or CD4 count (Eaton, Fraser, and HIV Portfolio). The Bendavid and Synthesis Transmission models simulate the change in viral load for infected individuals and associate HIV transmission with this according to an empirically described relationship [6]. Many models assume an increased risk of male-to-female transmission compared to female-to-male transmission, and attenuation in female-to-male transmission due to male circumcision. The Goals, STDSIM, and Synthesis Transmission models include an increased risk of HIV transmission in the presence of other sexually transmitted infections.

The models that simulate each individual's viral load (Bendavid and Synthesis Transmission) mechanistically relate reduction in transmission on treatment to the effect of ART on viral load, while the other models all assume a reduction in transmission of greater than 90% for individuals on ART. The Bendavid, Eaton, and Synthesis Transmission models simulate a period of a few months of incomplete viral suppression after ART initiation before the full reduction in infectiousness is achieved. These three models and EMOD include a return to higher infectiousness during treatment failure. The remaining models assume a fixed reduction in transmissibility as soon as treatment is started, until either death on ART or dropout from treatment. The Bendavid and Synthesis Transmission models simulate switching to second-line ART upon an immunologic (Bendavid) or virologic (Synthesis Transmission) failure event. The Synthesis Transmission model is the only model to explicitly simulate heterogeneous adherence to treatment between patients and the emergence and impact of resistance. The models vary in their assumptions about what happens to an individual after dropping out from treatment. The CD4 HIV/ART, Fraser, Goals, Granich, and HIV Portfolio models return individuals who drop out to an untreated state, allowing them to restart treatment in exactly the same manner as those that have never been treated, while the Bendavid, STDSIM, STI-HIV Interaction, and Synthesis Transmission models do not allow individuals to start treatment again in the implementation for this exercise. Eaton allows individuals to restart treatment, but only once, and EMOD allows half of individuals to restart treatment after they once again satisfy the eligibility criterion.

Eleven of the models simulate the South African national HIV epidemic, while the STDSIM model has been calibrated specifically to the higher prevalence Hlabisa subdistrict of KwaZulu-Natal Province, South Africa. Nine models were calibrated to reproduce the historical time series of HIV prevalence in South Africa, while the BBH, HIV Portfolio, and Bendavid models were initialized using the current epidemic state in the years 2009, 2011, and 2012, respectively, and simulated forward from that point.

Most of the models were calibrated to yield a single set of model parameters and outputs. Two of the models (Eaton and STI-HIV Interaction) were calibrated using a Bayesian framework allowing for uncertainty in model parameters, which produces a joint posterior distribution of parameter combinations consistent with the observed HIV epidemic [43,84]. The STI-HIV Interaction model allows for uncertainty in sexual behaviour, the natural history of HIV infection, and the effect of ART, while the Eaton model only allows for uncertainty in sexual behaviour and sexual mixing parameters.

Many of the models include facilities to simulate HIV testing and diagnosis, retention in care prior to treatment eligibility, and other processes related to achieving successful treatment, but these were not implemented for this exercise in order to conform to the simple intervention scenarios.

Table 2. Number of adults starting ART each year in the short term.

Year	"Low" Future Scale-Up	"Baseline" Future Scale-Up	"Medium" Future Scale-Up	"High" Future Scale-Up
2012	400,000	400,000	600,000	800,000
2013	200,000	400,000	600,000	900,000
2014	200,000	400,000	600,000	900,000
2015	200,000	400,000	600,000	700,000
2016	200,000	400,000	600,000	600,000
Total	1,200,000	2,000,000	3,000,000	3,900,000

Number of adults (age 15 y and older) initiating ART between midpoint of the previous year and the midpoint of indicated year.
doi:10.1371/journal.pmed.1001245.t002

Scenarios Representing the Existing ART Programme in South Africa

In a separate analysis, seven of the models (CD4 HIV/ART, Eaton, Fraser, Goals, Granich, STDSIM, and STI-HIV Interaction) were used to estimate the impact that the existing scale-up of ART in South Africa has had on HIV incidence and prevalence by comparing model simulations that include the ART scale-up over the past decade with the no-ART counterfactual. Models either used an existing calibration to the number of people on ART in South Africa (Fraser and STDSIM) or were calibrated using estimates of the number of adults starting and on ART in each year from 2001 to 2011 [45] (CD4 HIV/ART, Eaton, Goals, Granich, and STI-HIV Interaction).

Five models (Bendavid, CD4 HIV/ART, Eaton, Goals, and Granich) constructed short-term projections of HIV incidence in South Africa assuming different trajectories for continued ART scale-up from 2011 to 2016, the period covered by South Africa's national strategic plan [46]. Starting from the number of patients on ART in mid-2011, the numbers of adults starting ART in each of the years from mid-2011 through mid-2016 was specified. A "baseline" scenario was considered in which 400,000 adults would start ART in each of the next 5 y (approximately the number who started ART in 2009), for a total of 2 million new adults initiating ART. Three other scenarios were considered for the total numbers starting ART over the same period: (i) "low"—1.2 million start ART; (ii) "medium"—3 million start ART; and (iii) "high"—3.9 million start ART. (The exact number starting in each year is listed in Table 2.) The HIV incidence rate at the midpoint of 2016 and the cumulative number of new adult HIV infections over the period 2011 to 2016 were reported for each of these scenarios. For these projections, assumptions regarding CD4 distributions at ART initiation and rates of retention were based on actual treatment guidelines and programme experiences, but were not standardised across models.

Results

Figure 1 shows HIV prevalence and HIV incidence in 15- to 49-y-old males and females simulated by each of the models under the counterfactual assumption of no ART provision.

Other epidemiologic statistics are presented in Table 3. The estimates of adult male HIV prevalence in year 2012 ranged between 10% and 16%, and estimates of female prevalence between 17% and 23%. Male HIV incidence in year 2012 ranged between 1.1 and 2.0 per 100 person-years, and female incidence ranged between 1.7 and 2.6. The STDSIM model calibrated to KwaZulu-Natal Province simulated a considerably larger burden of HIV, consistent with observation [47], with prevalences in year 2012 of 23% and 33% in males and females, respectively, and incidence rates of 3.0 and 3.9 per 100 person-years, respectively. All of the sex-stratified models simulated higher HIV prevalence in adult women than in men, with sex ratios in HIV prevalence in year 2012 between 1.2 and 1.7, and all of the models except for Bendavid simulated higher incidence in year 2012 in females than in males.

Nearly all of the models projected declines in HIV incidence after 2012 in the absence of ART, but the magnitude of the projected natural changes between 2012 and 2050 varied widely from almost no change (Goals and Granich) to greater than 45% reduction (Bendavid and Synthesis Transmission).

Model projections of future national population growth in the absence of ART varied widely, ranging from a 6% reduction to a 13% increase in the population aged 15 y and older between the years 2012 and 2020. For comparison, the low and high variants for the projected total population growth from the United Nations Population Division over the same period (which incorporates some assumptions about ART provision) are 1.5% and 6.1% [48].

Impact of ART on HIV Incidence

Figure 2 presents the outcomes of an intervention starting in year 2012 with ART eligibility at CD4 count $\leq$350 cells/μl, reaching 80% of those requiring treatment, and retaining 85% of patients on ART after 3 y. This scenario reflects an optimised implementation of the current World Health Organization treatment guidelines [39] and the Joint United Nations Programme on HIV/AIDS definition for "universal access" of reaching 80% of those in need [32]. Compared to the no-treatment counterfactual scenario, ART provision reduced incidence in year 2020 by 35% to 54% across all models (Figure 2A). There was much greater variation, however, in the estimated long-term impact of the intervention. In year 2050, the range of the predicted reduction in incidence was from 32% to 74%. The relative impact of the ART intervention on HIV incidence decreased between 2020 and 2050 in four models and increased in seven.

Number of Person-Years of ART per Infection Averted

There was considerable variation between models in estimates of the number of person-years of treatment per infection averted. For the scenario described above, the range of estimates for the number of person-years of ART per infection averted between 2012 and 2020 was between 6.3 and 18.7, and over the period 2012 to 2050, the range was 4.5 to 20.2 (Figure 2B). The four models with the greatest estimates of the number of person-years of ART provided per infection averted (Eaton, EMOD, STI-HIV Interaction, and Synthesis Transmission) all explicitly included variation in transmission by age (e.g., allowing for reduced impact

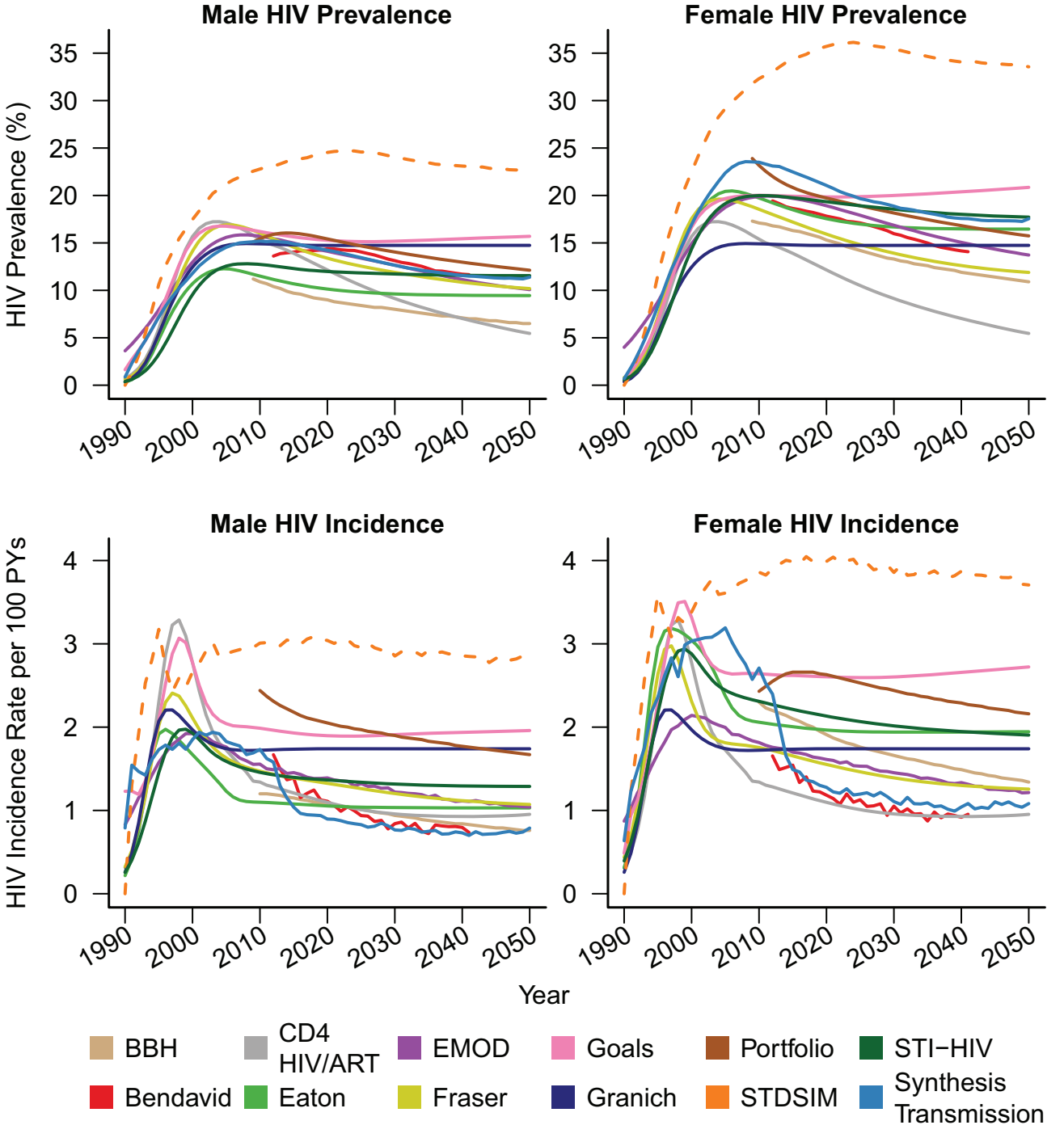

Figure 1. No-treatment counterfactual epidemic trends. Male (left) and female (right) HIV prevalence (top) and incidence (bottom) amongst 15- to 49-y-olds for counterfactual HIV epidemics with no ART. The STDSIM model is calibrated to a more severe epidemic in the Hlabisa subdistrict of KwaZulu-Natal Province, South Africa. The CD4 HIV/ART and Granich models do not stratify by sex, and the same prevalence and incidence curves are plotted for both sexes for these models. PYs, person-years.
doi:10.1371/journal.pmed.1001245.g001

through ART provision to older adults who are less sexually active and hence less likely to expose susceptible individuals), whereas the other models did not assume reduced transmission by older people. (STDSIM allows for decreased sexual activity for those older than 50 and has the lowest estimate of person-years of ART per infection averted, but simulates a much higher HIV incidence.)

Determinants of Programme Impact

The impact on incidence of increasing access from 50% to 100%, improving 3-y programme retention from 85% to 100%, and changing the CD4 threshold for treatment eligibility, is shown for each model in Figure 3. The reduction in incidence increases approximately linearly with access in all models. In

Table 3. Selected model outputs for counterfactual simulation with no ART.

Model Name	Age 15–49 y HIV Prevalence in Year 2012 (Percent)		Sex Ratio in Prevalence, Year 2012 (Female/Male)	Age 15–49 y HIV Incidence in Year 2012 (per 100 Person-Years)		Sex ratio in Incidence, Year 2012 (Female/Male)	Average Annual Population Growth Rate, Age 15+ y Population (per 100 People)		Year of Peak HIV Incidence		Percentage Change from Peak Incidence to year 2012 (Percent)		Percentage Change in Incidence, Year 2012 to 2020 (Percent)		Percentage Change in Incidence, Year 2020 to 2050 (Percent)	
	Male	Female		Male	Female		2012–2020	2012–2050	Male	Female	Male	Female	Male	Female	Male	Female
BBH	10.4	16.8	1.6	1.2	2.2	1.8	0.5	0.5					−9	−13	−30	−30
Bendavid	13.6	19.4	1.4	1.7	1.7	1.0	0.8	1.0[a]					−33	−29		
CD4 HIV/ART	14.8			1.3			1.5	1.1	1998		−61		−15		−13	
Eaton	11.1	19.2	1.7	1.1	2.0	1.9	1.4	1.5	1996	1997	−45	−36	−4	−4	−2	−1
EMOD	15.4	19.9	1.3	1.5	1.8	1.2	0.6	0.5	2001	2000	−23	−18	−7	−8	−25	−25
Fraser	15.4	18.0	1.2	1.4	1.7	1.2	−0.2	0.0	1997	1997	−40	−42	−8	−10	−19	−19
Goals	16.0	20.0	1.3	2.0	2.6	1.3	0.3	0.1	1998	1999	−36	−25	−3	−1	3	5
Granich	14.9			1.7			−0.2	−0.1	1997		−22		1		0	
HIV Portfolio	15.9	22.0	1.4	2.3	2.6	1.1	−0.8	−0.6					−11	3	−19	−18
STDSIM[b]	23.1	33.0	1.4	3.0	3.9	1.3	−1.3	−1.3	1995	2003	−6	3	2	2	−8	−7
STI-HIV	12.6	20.0	1.6	1.4	2.3	1.6	0.4	0.1	1999	1999	−28	−23	−5	−6	−6	−11
Synthesis Transmission	15.2	23.1	1.5	1.6	2.4	1.5	0.7	0.5	2001	2005	−18	−25	−43	−47	−12	−15

[a]Average annual growth rate for years 2012 to 2040.
[b]Model calibrated to HIV epidemic in KwaZulu-Natal Province.
doi:10.1371/journal.pmed.1001245.t003

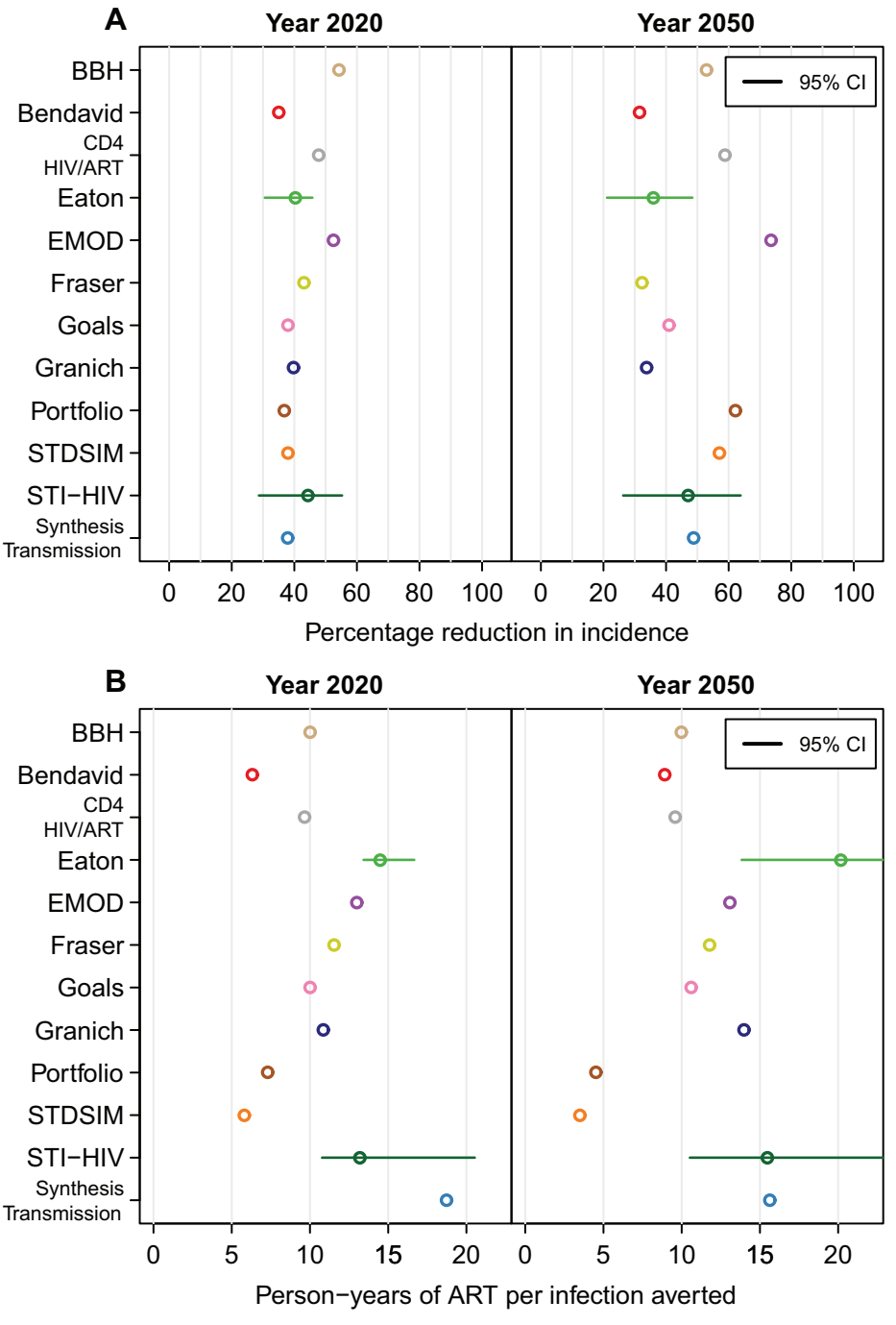

Figure 2. Impact of treatment for a scenario with eligibility at CD4≤350 cells/μl, 80% access, and 85% retention. (A) The percentage reduction in HIV incidence in the years 2020 and 2050 when eligibility for treatment is at CD4 count ≤350 cells/μl, 80% of individuals are treated, and 85% are retained on treatment after 3 y. (B) The cumulative number of person-years of ART provided per infection averted for the same scenario. Horizontal lines indicate 95% credible intervals (CI). For the Bendavid model, results in year 2040 are reported in the right panels.
doi:10.1371/journal.pmed.1001245.g002

most models, improvements in retention in care led to greater impact of treatment on HIV incidence. The benefit of improving retention was minimal for the Fraser, Granich, and HIV Portfolio models. Each of these models regards individuals who have dropped out of treatment identically to untreated eligible individuals, allowing them to start treatment again on average within 1 y. In several models, improved retention means that the impact improves more rapidly with increasing access (i.e., the slope in reduction in incidence as access increases is steeper for higher retention).

Figure 4 shows how the number of person-years of ART provided per infection averted up to year 2020 varied in relation to the intervention programme. There were no consistent trends across all models. In some models, with earlier initiation of treatment, fewer years of ART were required per infection averted (efficiency increases), whereas the opposite was predicted in others.

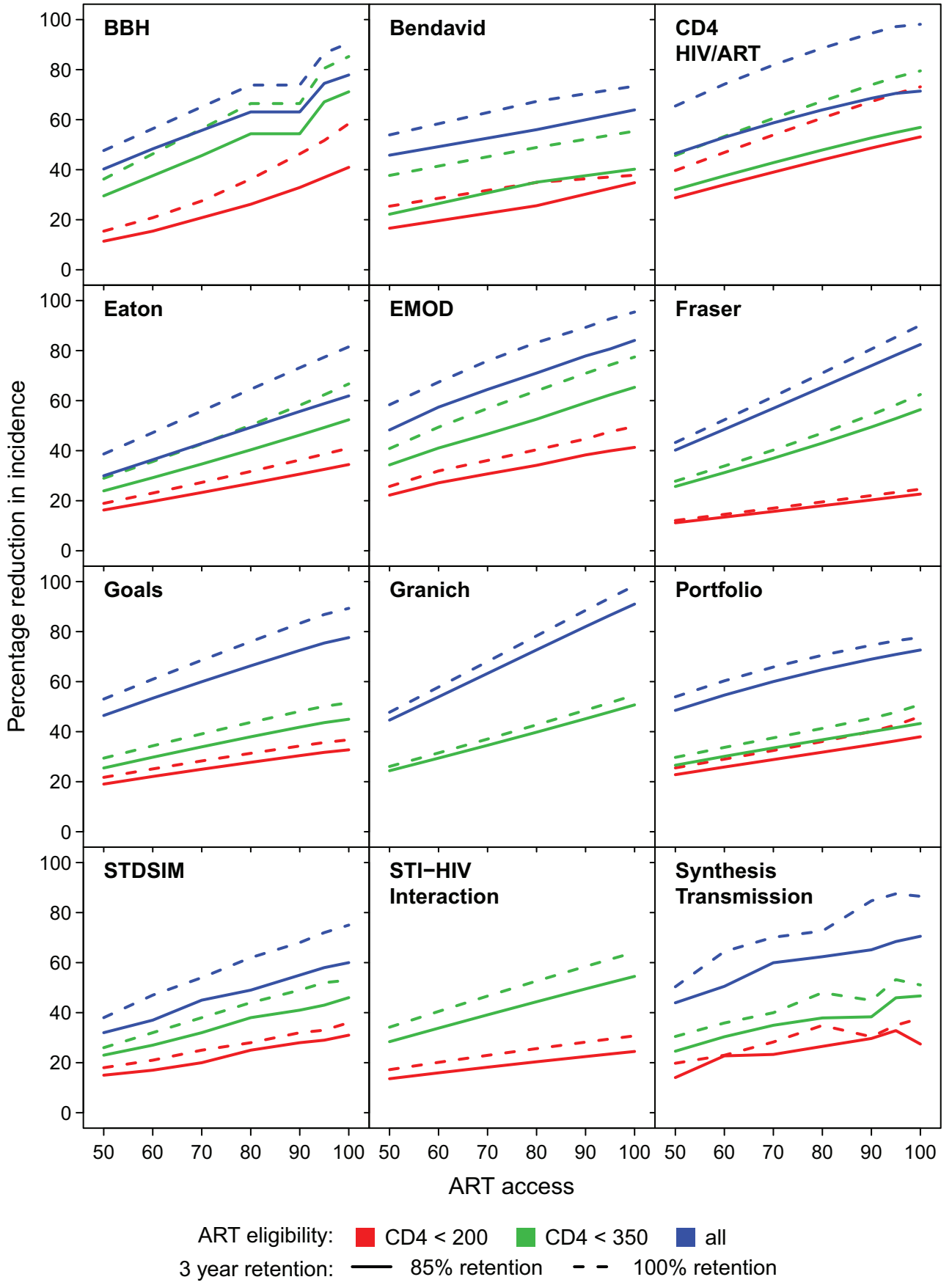
BBH
Bendavid
CD4 HIV/ART
Eaton
EMOD
Fraser
Goals
Granich
Portfolio
STDSIM
STI−HIV Interaction
Synthesis Transmission
Percentage reduction in incidence
0
20
40
60
80
100
50
60
70
80
90
100
ART access
ART eligibility: CD4 < 200 CD4 < 350 all
3 year retention: 85% retention 100% retention

Figure 3. Proportion reduction in HIV incidence in year 2020. For each model, the proportion reduction in HIV incidence in year 2020 for increasing access levels from 50% to 100% (horizontal axis). ART eligibility thresholds are indicated by line colour; 85% retention is indicated by solid lines, and perfect 100% retention is indicated by dashed lines.
doi:10.1371/journal.pmed.1001245.g003

For all of the models except the Granich model, which does not include increased transmissibility during late-stage infection, it might be expected that treating at lower CD4 count would be more efficient, as it targets treatment towards individuals with the highest current infectiousness (as in the BBH, Bendavid, CD4 HIV/ART, Eaton, and Goals models). However this could be counteracted if stage of infection interacts with other processes such as decreased propensity to form new partnerships with ageing. For half of the models (BBH, Eaton, EMOD, Fraser, Goals, Granich, and HIV Portfolio), increasing the percentage of the population with access to treatment reduced the amount of treatment per infection averted, at least at earlier CD4 initiation thresholds. This increased efficiency is indicative of increasing returns due to "herd immunity" at high intervention coverage levels.

Treatment Eligibility and the Theoretical Impact of "Test and Treat"

The models varied in their predictions as to the relative benefit of increasing treatment eligibility from a CD4 threshold of ≤200 cells/μl (national guidelines in some settings and close to actual experience in many) to ≤350 cells/μl (international guidelines) compared to further increasing eligibility to all infected individuals (Figure 3). The Bendavid, CD4 HIV/ART, Goals, HIV Portfolio, and Synthesis Transmission models all predicted that there would be only a relatively modest benefit in moving from initiation at ≤200 cells/μl to ≤350 cells/μl, and a much greater benefit in moving from initiation at ≤350 cells/μl to immediately upon diagnosis of HIV infection. In contrast, the BBH model simulated very little benefit in moving from the ≤350 cells/μl threshold to immediate eligibility. The Eaton, Fraser, and EMOD models showed similar benefits associated with each of the increments at moderate levels of access.

One important argument that has been made for immediate ART is that commitment of a large amount of ART now could reduce the cumulative amount of ART required in the future as a result of averted HIV infections [2,49]. Whether such savings could occur was evaluated by investigating whether the cumulative person-years of treatment through year 2050 to implement immediate treatment is less than the amount of ART required when treating after the CD4 count falls below 350 cells/μl for the same levels of access and retention. In six (BBH, CD4 HIV/ART, Fraser, Goals, HIV Portfolio, and STDSIM) out of eleven models (excluding STI-HIV Interaction) this was not the case: increasing eligibility from CD4≤350 cells/μl to immediate initiation always required more person-years of treatment, even with "perfect" ART programmes (100% access and 100% retention). However, for the EMOD model, expanding eligibility from CD4≤350 cells/μl to all HIV-infected adults required fewer cumulative person-years of treatment in all intervention scenarios (including access as low as 60% and retention in care as low as 75%). The Synthesis Transmission model found expanding access to be ART-saving with 70% access and retention above 95%, or with 80% access and retention above 85%. The other three models that found that expanding access could be ART-saving required more demanding assumptions about programmes: according to the Granich model, immediate initiation would be ART-saving if access were above 90% and retention above 95%; according to the Eaton model, access and retention would need to exceed 95%; and according to the Bendavid model, access and retention would both need to be 100%.

In an intervention treating all HIV-infected adults with 95% access and 95% retention, three (CD4 HIV/ART, EMOD, and HIV Portfolio) out of nine models (excluding BBH, Bendavid, and STI-HIV Interaction) predicted that HIV incidence would fall below 0.1% per year by 2050. The Granich model, which was used to argue the case for HIV elimination using treatment, projected that incidence in South Africa would be 0.13% under this scenario, a 92% reduction (in the original published projections, there was an assumption that risk of infection would fall by an additional 40% due to other interventions [34]).

Understanding Differences between Model Predictions

One factor expected to influence how much ART reduces HIV is the fraction of all transmission that occurs after individuals reach treatment eligibility thresholds, in the absence of any treatment [50]. Figure 5A shows the proportion of transmissions that occur from individuals in each CD4 count range in the counterfactual simulation in year 2012. Of the models that include a period of early infection, the percentage of new infections that occurs during this stage is between 4% and 28%, while between 20% and 51% of transmission results from individuals with CD4 cell count ≤200 cells/μl.

These percentages of transmission after ART eligibility can be compared with the percentage reduction in incidence in year 2020 (Figure 5B). Here, it is assumed that access is 80% and 3-y retention in care is 85%. Although this comparison explains why, within one model, earlier treatment initiation reduces HIV incidence more, the amount of between-model variation in projected impact explained by the distribution of transmission by CD4 count is modest. R^2 values for this relationship were 0.28, 0.20, and 0.40 for eligibility at CD4≤200, eligibility at CD4≤350, and immediate eligibility, respectively. The correlation did not improve when considering higher access or higher retention scenarios.

Two other factors hypothesized to explain the differences between the model projections are different assumptions about the efficacy of ART in reducing transmission—between 90% and 99%—and different assumptions about the outcomes of individuals who drop out from treatment programmes. To test the importance of these factors, selected intervention scenarios were repeated under the artificial assumption that an individual never transmits after initiating treatment (treatment is 100% efficacious at preventing transmission, and retention on treatment is 100%). This assumption increased the intervention impact in every model, but, surprisingly, did not reduce the variation in the results between models or improve the ability of factors such as different model assumptions about CD4 progression, HIV transmission, or the future trajectory of HIV incidence to explain the variation.

Estimates of the Current Impact of ART in South Africa

Figure 6 shows the estimated impact of the current ART programme in South Africa on HIV prevalence and incidence. The CD4 HIV/ART, Eaton, Goals, Granich, and STI-HIV Interaction models used estimates of the number of adults starting treatment in South Africa in each year between 2001 and 2011 from [45], and the Fraser and STDSIM models used

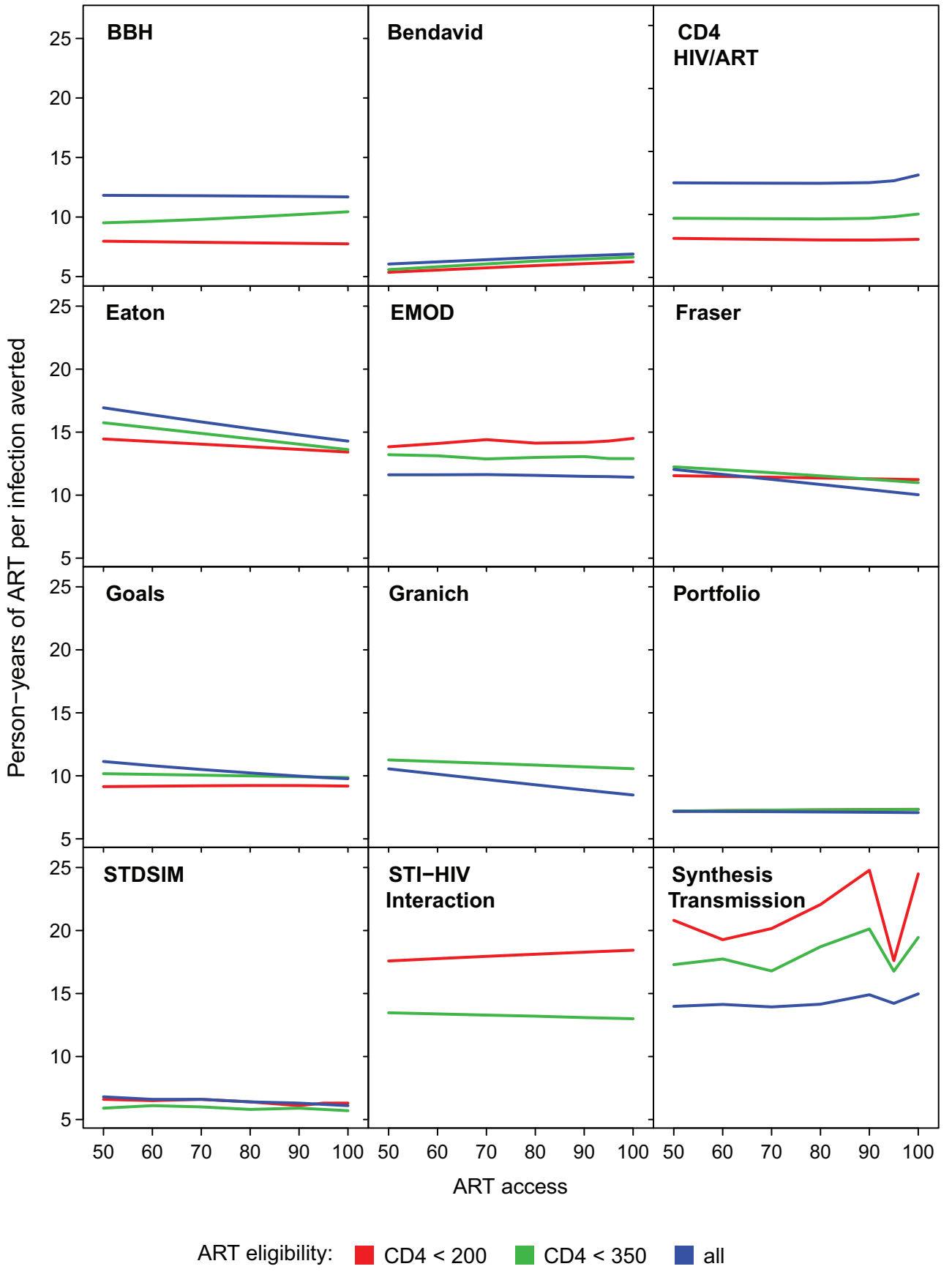
BBH
Bendavid
CD4
HIV/ART
Eaton
EMOD
Fraser
Goals
Granich
Portfolio
STDSIM
STI−HIV
Interaction
Synthesis
Transmission
Person−years of ART per infection averted
ART access
ART eligibility:
CD4 < 200
CD4 < 350
all

Figure 4. Cumulative number of person-years of ART provided per infection averted through year 2020. The cumulative person-years of ART provided per infection averted through the year 2020 for increasing access levels from 50% to 100% (horizontal axis), assuming 85% retention after 3 y. ART eligibility thresholds of are indicated by line colour. Varying retention did not affect trends between access and efficiency for any models.
doi:10.1371/journal.pmed.1001245.g004

existing calibrations to ART coverage levels in the Western Cape and KwaZulu-Natal Provinces, respectively. All of the models predicted that ART should already have had a substantial impact on the HIV epidemic, estimating that HIV incidence in year 2011 was between 17% and 32% lower than it would have been in the absence of ART. The increasing impact on HIV incidence over time mirrors the steep increase every year in the number of people starting treatment during this period.

The impact on prevalence was more modest and less consistent across models. The Eaton and STI-HIV Interaction models estimated that prevalence is around 8% higher than it would have been without treatment (an absolute increase in prevalence of one percentage point) due to the increased survival for those infected with HIV. The Fraser and Granich models suggest that this effect is offset by the reductions in incidence, so that there is no net change in prevalence. It is unlikely that standard surveillance methods based on monitoring trends in prevalence would have detected this impact, despite the significant underlying reductions in incidence.

The estimated potential impact of further ART scale-up is summarised in Table 4. In the baseline scenario, where 400,000 people are started on ART each year, the models estimated that incidence would be reduced in 2016 by between 13% and 26% compared to the incidence rate in 2011. If 800,000 fewer people are put on ART, then between 39,000 and 186,000 more new adult HIV infections would occur over the period 2012 to 2016 than under the baseline scenario. If more people are put on ART—3.0 or 3.9 million over the next 5 y—then the models estimated that the number of new infections over the 5-y period would be reduced by 64,000 to 327,000 and 270,000 to 521,000, respectively, compared to the baseline. The table underscores that there are still substantial potential preventive benefits from expanding ART coverage in South Africa. The models that tended to estimate the greatest reduction in incidence in hypothetical programmes over the medium term (CD4 HIV/ART, Goals, and Granich) also tended to project greater reductions in incidence over the short term in these more realistic scenarios.

Discussion

The mathematical models used to simulate the impact of treatment on HIV incidence in South Africa are diverse in their structure, level of complexity, representation of the HIV transmission process and the ART intervention, and parameter choices. All twelve of the models compared in this analysis predicted that treatment could substantially reduce HIV incidence—even using past or existing treatment guideline eligibility criteria, provided that coverage is high. Only three (CD4 HIV/ART, EMOD, and HIV Portfolio) out of nine models (excluding BBH, Bendavid, and STI-HIV Interaction), however, predicted that treatment could reduce HIV incidence below 0.1% by year 2050 (the definition of "elimination" established by [34]), even with very high access and

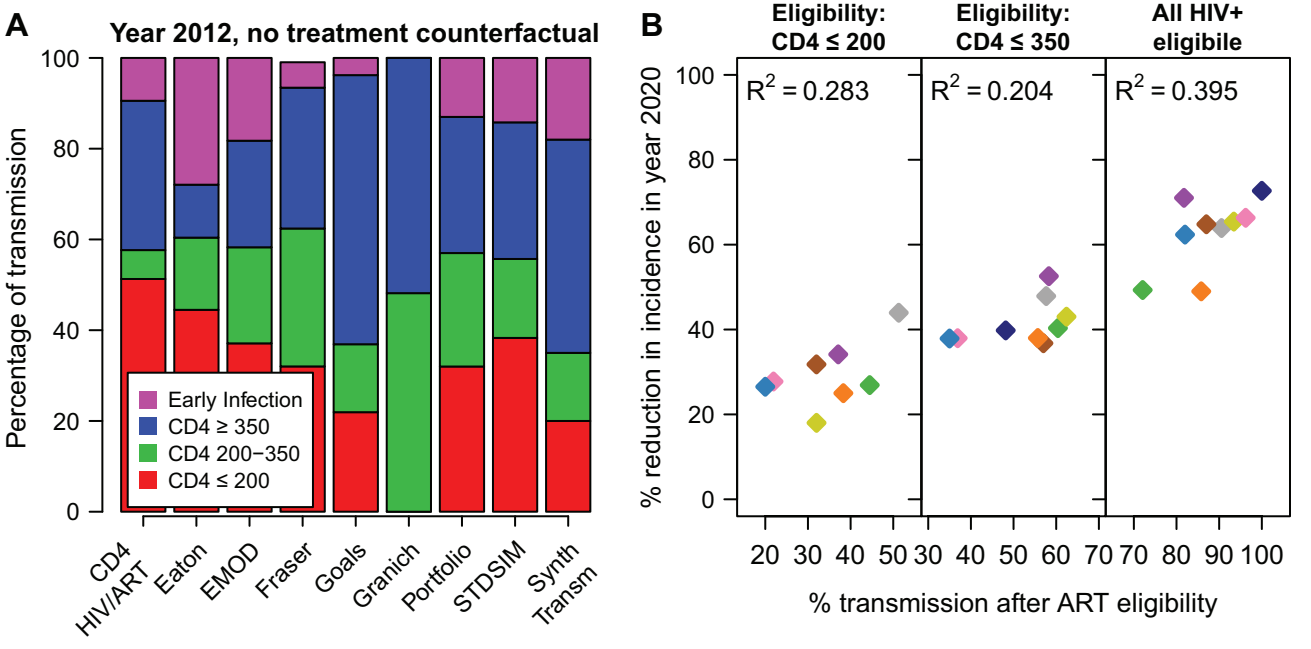

Figure 5. Impact of treatment by transmission in each CD4 category. (A) The percentage of all HIV transmissions from individuals in each CD4 cell count category in year 2012, in the no-ART counterfactual simulation. (B) The reduction in incidence in year 2020, for the 80% access and 85% retention scenario, according to the cumulative proportion of transmission that occurs after eligibility (A). For the scenario where all HIV-positive adults are eligible ("all HIV+ eligible"), the percentage of transmission after ART eligibility is the percentage of transmission that occurs after the end of primary HIV infection. Colours for models are the same as in Figures 1 and 2. The BBH, Bendavid, and STI-HIV Interaction models do not estimate the proportion of transmission in each CD4 category and are not included in this figure.
doi:10.1371/journal.pmed.1001245.g005

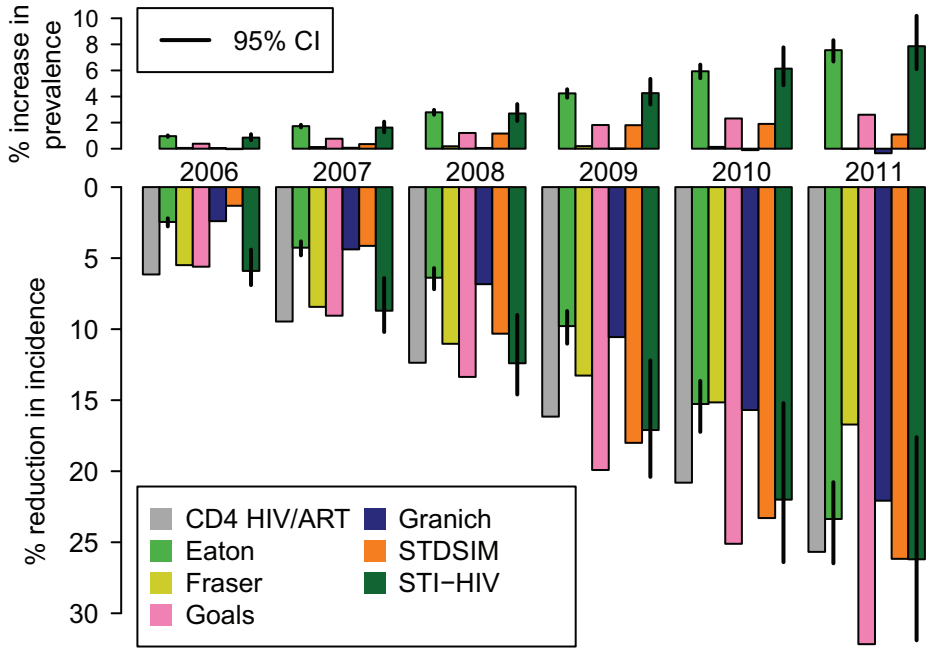

Figure 6. The impact of the existing ART programme in South Africa on HIV prevalence and incidence. The percentage increase in HIV prevalence (top) and the percentage reduction in HIV incidence rate (bottom) compared to what would have occurred in the absence of any ART for years the 2006 to 2011. These are estimated by comparing HIV prevalence and incidence in a model calibrated to the existing scale-up of ART in South Africa from 2001 to 2011 with a model simulation with no ART provision. The CD4 HIV/ART, Eaton, Goals, Granich, and STI-HIV Interaction models use the same estimates of the number starting ART each year (from [45]). Fraser uses an existing calibration to the ART scale-up in the Western Cape Province. STDSIM is calibrated to the number of people on ART in the Hlabisa subdistrict. Vertical lines on the Eaton and STI-HIV Interaction models indicate 95% credible intervals (CI).
doi:10.1371/journal.pmed.1001245.g006

retention. When simulating the historical scale-up of ART in South Africa, the models indicated that ART may already have reduced HIV incidence by between 17% and 32% in 2011, compared to what would have been expected in the absence of ART.

Although there have been ad hoc informal model comparison exercises [51], collections of work using standardised assumptions for interventions [52], and thorough model comparisons involving a few research groups [53,54], to our knowledge, this exercise is the first to bring together such a large number of independent modelling groups to examine the same set of interventions. We hope that this will provide a foundation for much more collaborative work.

In this study we set out to test whether different models of the potential impact of treatment on new HIV infections in South Africa would make similar predictions when implementing the same intervention scenarios. We found substantial consistency between the model projections of the impact of ART interventions on HIV incidence in the short term (8 y). However, there was more variation in the predicted longer term (38 y) reductions in incidence, and models also produced divergent estimates of the number of person-years of ART provided per infection averted. While establishing where models agree and disagree about the epidemiological impact of ART represents an important scientific finding in itself, the substantial variation in the long-term impact and efficiency of interventions demands further investigation and explanation.

Based on epidemiological theory and previous modelling studies, we hypothesized a number of model attributes that might explain differences in model predictions about the impact of ART, including the amount of transmission in different stages of HIV infection, the assumed efficacy of ART for preventing transmission, opportunities for treatment reinitiation following dropout from a treatment programme, the age and sex structure of the population, future population growth rates, the degree of heterogeneity and assortativity in sexual mixing, the future trajectory of HIV incidence in the absence of intervention, and the inclusion of changes in sexual behaviour over the past decade. There was indeed substantial variation between the models in their characterisation of each of these aspects of the system, largely reflecting the true uncertainties that persist even after decades of tremendous research into the epidemiology of HIV in South Africa. We were able to show that crude differences in the proportion of transmission at each stage of infection explained a modest amount of the variation in the short-term impact of ART, but less of the long-term impact. However, beyond this, findings from the models did not appear to clearly support any of these hypotheses in univariate analyses, likely because of the large number of processes that interact nonlinearly to create HIV epidemics and interventions. For example, projecting a seemingly simple quantity such as the number of person-years of ART that will be provided in an intervention depends on future population growth, the natural trend in the epidemic, the proportion of HIV-infected individuals qualifying for treatment, retention and survival on ART, and the impact that ART provision has on future HIV incidence. This situation contrasts with that of an earlier exercise that compared predictions of the impact of male

Table 4. HIV incidence rate per 100 person-years in year 2016 for different potential scenarios of future ART scale-up.

Model	"Low" Future Scale-Up		"Baseline" Future Scale-Up		"Medium" Future Scale-Up		"High" Future Scale-Up	
	HIV Incidence 2016[a]	Number Different from "Baseline"[,b]	HIV Incidence 2016[a]	Number Different from "Baseline"[,b]	HIV Incidence 2016[a]	Number Different from "Baseline"[,b]	HIV Incidence 2016[a]	Number Different from "Baseline"[,b]
Bendavid	1.20	39,000	1.14	—	1.06	−64,000	0.89	−270,000
CD4 HIV/ART	1.77	176,000	1.21	—	0.80	−327,000	0.54	−506,000
Eaton[c]	1.23 (1.07, 1.39)	106,000 (100,000, 113,000)	1.05 (0.89, 1.21)	—	0.79 (0.65, 0.95)	−111,000 (−118,000, −101,000)	0.70 (0.56, 0.83)	−310,000 (−334,000, −286,000)
Goals	1.50	186,000	1.16	—	0.88	−187,000	0.64	−419,000
Granich	1.34	142,000	1.14	—	0.87	−225,000	0.53	−521,000

Number of adults (age 15 y and older) initiating ART between midpoint of the previous year and the midpoint of indicated year.

[a]HIV incidence rate per 100 susceptible person-years amongst 15- to 49-y-olds at midpoint of year 2016.

[b]Cumulative number of additional new infections over the period mid-2011 to mid-2016 compared to "baseline" future scale-up scenario (rounded to nearest 1,000).

[c]The Eaton model reports posterior mean and 95% credible intervals.

doi:10.1371/journal.pmed.1001245.t004

circumcision interventions [51], where the relationship between the established efficacy of the intervention and population-level impact was less complicated.

Having investigated the extent to which differences in ART programmes determine differences in results, the natural focus for future model comparison studies should be to explore the contribution of other hypotheses through incrementally standardising biological, behavioural, and demographic model parameters, and calibrating models to the same levels of HIV prevalence and incidence. A systematic approach to standardising model parameters would identify which parameters most significantly influence the results and guide priorities for future data collection. The HIV Modelling Consortium (http://www.hivmodelling.org) will coordinate such research efforts in coming months to investigate the extent to which variation in model predictions is driven by differences in underlying models of sexual mixing, or different models of the natural history of infection and epidemic trajectory.

Although our experiment and analysis has focused on how factors included in models can affect model predictions, it is important to note that if all models exclude an important aspect of the system, they could all be wrong. Early models of the impact of ART on HIV incidence were very focused on the concern that increased sexual risk behaviour might offset the reduction in transmission for those on treatment, but for this exercise all of the models assumed that population risk behaviour would not change in response to the introduction of ART. This may be a reasonable assumption given consistent evidence that patients report safer sexual behaviour after starting ART [55–59] and given the relative lack of information from sub-Saharan Africa about how the untreated and HIV-negative population responds to the availability of treatment [60]. But in other epidemic settings the availability of ART has been associated with receding gains in protective behaviour [61–63], and monitoring this in sub-Saharan African settings will be a priority for surveillance over coming years. The models also all assumed high efficacy of ART to reduce transmission. True effectiveness will depend on adherence and the level of viral suppression, which is mainly determined by adherence levels. While there are some data from South Africa on viral suppression rates outside carefully controlled trial settings [64], further information on this and on patterns of acquired and transmitted resistance will help in the calibration of models. Only one of the models in this exercise (Synthesis Transmission) explicitly incorporated the effect of antiretroviral drug resistance on the impact of ART interventions. Models have predicted that antiretroviral drug resistance could be widespread in sub-Saharan Africa in coming decades [20], which could eventually lead to the spread of transmitted drug resistance [65,66]. This could affect the long-term costs and efficacy of treatment-as-prevention strategies [67].

Another finding from systematically comparing models is that often seemingly independent modelling studies rely on the same limited data. Nearly all of the models relied on two sources to derive parameters for elevated infectiousness during the first few weeks of infection [68,69], but both of these sources are based principally on data from a few retrospective couples in Rakai, Uganda [70] (see [71]). This highlights both how invaluable these data are and also the importance of recognising the dependencies between seemingly independent modelling studies. However, even using the same data, models may reach different conclusions. The Eaton, EMOD, and Fraser models all in some way used the estimates of early HIV infectivity from [69] but estimate very different contributions of this stage to overall HIV transmission (Figure 4A), and the three models all reached different conclusions

from those in another recent modelling study relying on these same estimates [72].

The purpose of this exercise was not to draw conclusions or recommendations about specific ART intervention strategies, but rather to test the hypothesis that a range of different models would come to similar conclusions about the impact of ART on HIV incidence when the same interventions were modelled. The simulated interventions were artificially simple and stylized to enable comparison between models. These did not explicitly simulate the steps of HIV testing, diagnosis, linkage to care, and adherence to ART required to achieve the access levels specified in the intervention scenarios (although several of the models include facility for this and have investigated this in independent analyses). Interpretation of models simulating high levels of treatment coverage should be cautioned by data suggesting that at present fewer than one-third of patients in sub-Saharan Africa are continuously retained in care from HIV diagnosis to ART initiation [73], and that barriers remain to access to and uptake of HIV testing [32]. The models assumed that all individuals eligible for treatment were equally likely to access treatment, which might not be true in practice (for example, women are more likely to start treatment than men [74]). The comparison scenario (counterfactual) against which interventions were evaluated assumed no treatment at all, which made it easier to compare models, but is clearly not the relevant benchmark for policy-makers. This study has also considered treatment in isolation from other interventions, even as there is broad consensus that "combination prevention" strategies are presently the best strategy for attacking the epidemic [41,75].

We hope that this study will help to characterise the models that are being used to investigate questions related to the impact of HIV treatment and enable those who rely on models for decision-making to think critically about how the assumptions underlying models affect the results. The relative consistency between models' estimates of the short-term epidemiological impact of ART, including the impact of the existing ART programme, provides some reassurance that model projections on this time scale may be relatively robust to the substantial uncertainties in parameters and systems. This is a significant result considering that such short-term projections are often the most relevant for policy and resource allocation questions. On the other hand, the substantial variation in long-term epidemiological impacts and efficiency of ART, upon which arguments of substantial epidemic reduction and cost savings hinge, suggests that results in these areas from any single model should be extrapolated with caution. Care should be taken to ensure that models evaluating the long-term costs, benefits, and cost-effectiveness of treatment programmes adequately communicate the degree and myriad sources of uncertainty that influence these outputs.

A common question when faced with a diversity of model results is whether some models are "better" or "worse". Without data against which to test the predictions of models, it is not possible to answer this question in a study such as this, nor is this the correct question to be asking. Rather, users of model outputs should ask whether models include the necessary components to capably answer the specific questions at hand, and whether the models make credible assumptions in light of the information available, and choose models accordingly. Evaluated along these guidelines, the most appropriate models will vary between applications, so there is no single "best" model. However, in this exercise, the models that tended to project more "pessimistic" outcomes for the interventions seemed to do so for important reasons. For example, models that estimated poorer efficiency of ART for averting infections tended to be those that simulated ART provision for those at older ages, who might be at lower risk of transmitting, or included the elevated risk of transmission for those failing treatment, whereas models with more optimistic predictions assumed that risk behaviour did not vary by age or that transmission was fully suppressed immediately upon beginning treatment until death on ART or dropout. Artificial convergence of models should be avoided when true uncertainties persist about the system. It is incumbent upon modellers to incorporate and communicate uncertainty in projections, and identify which components of the system account for the uncertainty. For this exercise, only one model (STI-HIV Interaction) included a comprehensive analysis accounting for uncertainty about basic epidemiology and intervention efficacy. While the focus of the study was on variation between models, it is interesting to observe that the 95% credible interval representing parameter uncertainty for this model encompassed the point estimates of the other eleven models.

Fortunately there will be important new opportunities in the near future to test, validate, and improve epidemiological models of HIV treatment. These include comparing projections to the experience of expanded ART in industrialised countries [61,63], the observed impact of ART in well-characterised communities [76], and results of a number of community-randomized trials of treatment as prevention that will soon be underway [44]. As new data are reported, the accuracy of models projecting the impact of treatment as prevention should improve, and we expect that validated and scientifically based model projections will continue to be central in understanding how ART can have the greatest impact in mitigating the global HIV epidemic.

Acknowledgments

Results of this investigation were first presented at a meeting of the HIV Modelling Consortium hosted by the South African Centre for Epidemiological Modelling and Analysis (SACEMA), 3–5 November 2011, in Stellenbosch, South Africa. The authors thank all participants in this meeting for feedback and insights into the results of this study (see [49] for participant list). We would like to thank Britta Jewell for her assistance with facilitating this study.

V. C. and A. N. P. acknowledge valuable input into model development of the Synthesis Transmission model from Jens Lundgren, Deenan Pillay, and Geoff Garnett, and thank ISD Research Computing, University College London, for the use of the Legion distributed memory computing cluster. C. F. thanks Anne Cori for assistance with preparing the description of the model.

Author Contributions

Conceived and designed the experiments: JWE JAS BGW TBH. Performed the experiments: JWE LFJ TB EB AB DEB VC CF JACH SH DJK EFL ANP CP JS EW. Analyzed the data: JWE. Wrote the first draft of the manuscript: JWE. Contributed to the writing of the manuscript: JWE LFJ JAS TB EB AB DEB VC CF JACH SH DJK EFL ANP CP JS EAW BGW TBH. ICMJE criteria for authorship read and met: JWE LFJ JAS TB EB AB DEB VC CF JACH SH DJK EFL ANP CP JS EAW BGW TBH. Agree with manuscript results and conclusions: JWE LFJ JAS TB EB AB DEB VC CF JACH SH DJK EFL AP CP JS EAW BGW TBH.

References

1. Montaner JSG, Hogg R, Wood E, Kerr T, Tyndall M, et al. (2006) The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 368: 531–536. doi:10.1016/S0140-6736(06)69162-9

2. De Cock KM, Gilks CF, Lo Y-R, Guerma T (2009) Can antiretroviral therapy eliminate HIV transmission? Lancet 373: 7–9. doi:10.1016/S0140-6736(08)61732-8

3. Garnett GP, Baggaley RF (2009) Treating our way out of the HIV pandemic: could we, would we, should we? Lancet 373: 9–11. doi:10.1016/S0140-6736(08)61698-0
4. Cohen J (2011) Halting HIV/AIDS epidemics. Science 334: 1338–1340.
5. Granich R, Gupta S, Suthar A, Smyth C, Hoos D, et al. (2011) Antiretroviral therapy in prevention of HIV and TB: update on current research efforts. Curr HIV Res 9: 446–469. doi:10.2174/157016211798038597
6. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, et al. (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. N Engl J Med 342: 921–929. doi:10.1056/NEJM200003303421303
7. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP (2007) Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. Proc Natl Acad Sci U S A 104: 17441–17446. doi:10.1073/pnas.0708559104
8. Vernazza PL, Troiani L, Flepp MJ, Cone RW, Schock J, et al. (2000) Potent antiretroviral treatment of HIV-infection results in suppression of the seminal shedding of HIV. The Swiss HIV Cohort Study. AIDS 14: 117–121.
9. Cu-Uvin S, Caliendo AM, Reinert S, Chang A, Juliano-Remollino C, et al. (2000) Effect of highly active antiretroviral therapy on cervicovaginal HIV-1 RNA. AIDS 14: 415–421.
10. Baggaley RF, Ferguson NM, Garnett GP (2005) The epidemiological impact of antiretroviral use predicted by mathematical models: a review. Emerg Themes Epidemiol 2: 9. doi:10.1186/1742-7622-2-9
11. Dodd PJ, Garnett GP, Hallett TB (2010) Examining the promise of HIV elimination by "test and treat" in hyperendemic settings. AIDS 24: 729–735. doi:10.1097/QAD.0b013e32833433fc
12. Anderson RM, Gupta S, May RM (1991) Potential of community-wide chemotherapy or immunotherapy to control the spread of HIV-1. Nature 350: 356–359. doi:10.1038/350356a0
13. Gilliam BL, Dyer JR, Fiscus SA, Marcus C, Zhou S, et al. (1997) Effects of reverse transcriptase inhibitor therapy on the HIV-1 viral burden in semen. J Acquir Immune Defic Syndr Hum Retrovirol 15: 54–60.
14. Blower SM, Gershengorn HB, Grant RM (2000) A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 287: 650–654.
15. Dangerfield BC, Fang Y, Roberts CA (2001) Model-based scenarios for the epidemiology of HIV/AIDS: the consequences of highly active antiretroviral therapy. Syst Dyn Rev 17: 119–150. doi:10.1002/sdr.211
16. Law MG, Prestage G, Grulich A, Van de Ven P, Kippax S (2001) Modelling the effect of combination antiretroviral treatments on HIV incidence. AIDS 15: 1287–1294.
17. Velasco-Hernandez JX, Gershengorn HB, Blower SM (2002) Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? Lancet Infect Dis 2: 487–493. doi:10.1016/S1473-3099(02)00346-8
18. Xiridou M, Geskus R, De Wit J, Coutinho R, Kretzschmar M (2003) The contribution of steady and casual partnerships to the incidence of HIV infection among homosexual men in Amsterdam. AIDS 17: 1029–1038. doi:10.1097/01.aids.0000050879.72891.1d
19. Gray RH, Li X, Wawer MJ, Gange SJ, Serwadda D, et al. (2003) Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda. AIDS 17: 1941–1951. doi:10.1097/01.aids.0000076313.76477.fc
20. Blower S, Bodine E, Kahn J, McFarland W (2005) The antiretroviral rollout and drug-resistant HIV in Africa: insights from empirical data and theoretical models. AIDS 19: 1–14.
21. Baggaley RF, Garnett GP, Ferguson NM (2006) Modelling the impact of antiretroviral use in resource-poor settings. PLoS Med 3: e124. doi:10.1371/journal.pmed.0030124
22. McCormick AW, Walensky RP, Lipsitch M, Losina E, Hsu H, et al. (2007) The effect of antiretroviral therapy on secondary transmission of HIV among men who have sex with men. Clin Infect Dis 44: 1115–1122. doi:10.1086/512816
23. Salomon JA, Hogan DR (2008) Evaluating the impact of antiretroviral therapy on HIV transmission. AIDS 22(Suppl 1): S149–S159. doi:10.1097/01.aids.0000327636.82542.87
24. Attia S, Egger M, Müller M, Zwahlen M, Low N (2009) Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. AIDS 23: 1397–1404. doi:10.1097/QAD.0b013e32832b7dca
25. Reynolds SJ, Makumbi F, Nakigozi G, Kagaayi J, Gray RH, et al. (2011) HIV-1 transmission among HIV-1 discordant couples before and after the introduction of antiretroviral therapy. AIDS 25: 473–477. doi:10.1097/QAD.0b013e3283437c2b
26. Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, et al. (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet 375: 2092–2098. doi:10.1016/S0140-6736(10)60705-2
27. Glynn JR, Price A, Floyd S, Molesworth A, Kayuni N, et al. (2011) Anitretroviral therapy reduces HIV transmission in discordant couples in northern Malawi. Sex Transm Infect 87: A218. doi:10.1136/sextrans-2011-050108.277
28. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, et al. (2011) Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 365: 493–505. doi:10.1056/NEJMoa1105243
29. Lodwick R, Costagliola D, Reiss P, Torti C, Teira R, et al. (2010) Triple-class virologic failure in HIV-infected patients undergoing antiretroviral therapy for up to 10 years. Arch Intern Med 170: 410–419. doi:10.1001/archinternmed.2009.472
30. Boulle A, Van Cutsem G, Hilderbrand K, Cragg C, Abrahams M, et al. (2010) Seven-year experience of a primary care antiretroviral treatment programme in Khayelitsha, South Africa. AIDS 24: 563–572. doi:10.1097/QAD.0b013e328333bfb7
31. Nash D, Katyal M, Brinkhof MWG, Keiser O, May M, et al. (2008) Long-term immunologic response to antiretroviral therapy in low-income countries: a collaborative analysis of prospective studies. AIDS 22: 2291–2302. doi:10.1097/QAD.0b013e3283121ca9
32. Joint United Nations Programme on HIV/AIDS (2010) Towards universal access: scaling up priority HIV/AIDS interventions in the health sector. Progress report 2010. Available: http://whqlibdoc.who.int/publications/2010/9789241500395_eng.pdf. Accessed 19 December 2011.
33. Joint United Nations Programme on HIV/AIDS (2011) World AIDS Day Report 2011. How to get to zero: faster, smarter, better. Available: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216_WorldAIDSday_report_2011_en.pdf. Accessed 23 December 2011.
34. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57. doi:10.1016/S0140-6736(08)61697-9
35. Wagner BG, Blower S (2009) Voluntary universal testing and treatment is unlikely to lead to HIV elimination: a modeling analysis. doi:10.1038/npre.2009.3917.1
36. Kretzschmar ME, van der Loeff MF, Coutinho RA (2012) Elimination of HIV by test and treat. AIDS 26: 247–248. doi:10.1097/QAD.0b013e32834e1592
37. Bacaër N, Pretorius C, Auvert B (2010) An age-structured model for the potential impact of generalized access to antiretrovirals on the South African HIV epidemic. Bull Math Biol 72: 2180–2198. doi:10.1007/s11538-010-9535-2
38. Bendavid E, Brandeau ML, Wood R, Owens DK (2010) Comparative effectiveness of HIV testing and treatment in highly endemic regions. Arch Intern Med 170: 1347–1354. doi:10.1001/archinternmed.2010.249
39. World Health Organization (2010) Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach, 2010 revision. Available: http://whqlibdoc.who.int/publications/2010/9789241599764_eng.pdf. Accessed 22 December 2011.
40. PEPFAR Scientific Advisory Board (2011) PEPFAR Scientific Advisory Board recommendations for the Office of the US Global AIDS Coordinator: implications of HPTN 052 for PEPFAR's treatment programs. Available: http://www.pepfar.gov/documents/organization/177126.pdf. Accessed 19 December 2011.
41. Schwartländer B, Stover J, Hallett T, Atun R, Avila C, et al. (2011) Towards an improved investment approach for an effective response to HIV/AIDS. Lancet 377: 2031–2041. doi:10.1016/S0140-6736(11)60702-2
42. Hontelez JA, de Vlas SJ, Tanser F, Bakker R, Bärnighausen T, et al. (2011) The impact of the new WHO antiretroviral treatment guidelines on HIV epidemic dynamics and cost in South Africa. PLoS ONE 6: e21919. doi:10.1371/journal.pone.0021919
43. Johnson LF, Hallett TB, Rehle TM, Dorrington RE (2012) The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. J R Soc Interface 9: 1544–1554. doi:10.1098/rsif.2011.0826
44. Boily MC, Mâsse B, Alsallaq R, Padian NS, Eaton JW, et al. (2012) HIV treatment as prevention: considerations in the design, conduct, and analysis of cluster randomized controlled trials of combination HIV prevention. PLoS Med 9: e1001250. doi:10.1371/journal.pmed.1001250
45. Johnson LF (2012) Access to antiretroviral treatment in South Africa, 2004–2011. South Afr J HIV Med 13: 22–27.
46. South Africa National AIDS Council (2011) National strategic plan on HIV, STIs and TB: 2012–2016 summary. Available: http://www.doh.gov.za/docs/stratdocs/2012/NSPsum.pdf. Accessed 25 May 2012.
47. Shisana O, Rehle T, Simbayi LC, Zuma K, Jooste S, et al. (2009) South African national HIV prevalence, incidence, behaviour and communication survey 2008: a turning tide amongst teenagers? Cape Town: HSRC Press.
48. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (2011) World population prospects: the 2010 revision. Available: http://esa.un.org/unpd/wpp/index.htm. Accessed 19 December 2011.
49. The HIV Modelling Consortium Treatment as Prevention Editorial Writing Group (2012) HIV treatment as prevention: models, data, and questions—Towards evidence-based decision-making. PLoS Med 9: e1001259. doi:10.1371/journal.pmed.1001259
50. Anderson RM, May RM (1992) Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press.
51. UNAIDS/WHO/SACEMA Expert Group on Modelling the Impact and Cost of Male Circumcision for HIV Prevention (2009) Male circumcision for HIV prevention in high HIV prevalence settings: what can mathematical modelling contribute to informed decision making? PLoS Med 6: e1000109. doi:10.1371/journal.pmed.1000109
52. Hankins CA, Glasser JW, Chen RT (2011) Modeling the impact of RV144-like vaccines on HIV transmission. Vaccine 29: 6069–6071. doi:10.1016/j.vaccine.2011.07.001

53. Halloran ME, Ferguson NM, Eubank S, Longini IM, Cummings DA, et al. (2008) Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci U S A 105: 4639–4644. doi:10.1073/pnas.0706849105
54. Althaus CL, Turner KM, Schmid BV, Heijne JC, Kretzschmar M, et al. (2011) Transmission of Chlamydia trachomatis through sexual partnerships: a comparison between three individual-based models and empirical data. J R Soc Interface 9: 136–146. doi:10.1098/rsif.2011.0131
55. Bateganya M, Colfax G, Shafer LA, Kityo C, Mugyenyi P, et al. (2005) Antiretroviral therapy and sexual behavior: a comparative study between antiretroviral-naive and -experienced patients at an urban HIV/AIDS care and research center in Kampala, Uganda. AIDS Patient Care STDS 19: 760–768.
56. Bunnell R, Ekwaru JP, Solberg P, Wamai N, Bikaako-Kajura W, et al. (2006) Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda. AIDS 20: 85–92.
57. Eisele TP, Mathews C, Chopra M, Lurie MN, Brown L, et al. (2009) Changes in risk behavior among HIV-positive patients during their first year of antiretroviral therapy in Cape Town South Africa. AIDS Behav 13: 1097–1105. doi:10.1007/s10461-008-9473-2
58. McClelland RS, Graham SM, Richardson BA, Peshu N, Masese LN, et al. (2010) Treatment with antiretroviral therapy is not associated with increased sexual risk behavior in Kenyan female sex workers. AIDS 24: 891–897. doi:10.1097/QAD.0b013e32833616c7
59. Venkatesh KK, de Bruyn G, Lurie MN, Mohapi L, Pronyk P, et al. (2010) Decreased sexual risk behavior in the era of HAART among HIV-infected urban and rural South Africans attending primary care clinics. AIDS 24: 2687–2696. doi:10.1097/QAD.0b013e32833e78d4
60. Venkatesh KK, Flanigan TP, Mayer KH (2011) Is expanded HIV treatment preventing new infections? Impact of antiretroviral therapy on sexual risk behaviors in the developing world. AIDS 25: 1939–1949. doi:10.1097/QAD.0b013e32834b4ced
61. Bezemer D, de Wolf F, Boerlijst MC, van Sighem A, Hollingsworth TD, et al. (2008) A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. AIDS 22: 1071–1077. doi:10.1097/QAD.0b013e3282fd167c
62. Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS (2012) HIV treatment as prevention: the utility and limitations of ecological observation. PLoS Med 9: e1001260. doi:10.1371/journal.pmed.1001260
63. Wilson DP (2012) HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med 9: e1001231. doi:10.1371/journal.pmed.1001231
64. Barth RE, van der Loeff MF, Schuurman R, Hoepelman AI, Wensing AM (2010) Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa: a systematic review. Lancet Infect Dis 10: 155–166. doi:10.1016/S1473-3099(09)70328-7
65. Vardavas R, Blower S (2007) The emergence of HIV transmitted resistance in Botswana: "when will the WHO detection threshold be exceeded?". PLoS ONE 2: e152. doi:10.1371/journal.pone.0000152
66. Phillips AN, Pillay D, Garnett G, Bennett D, Vitoria M, et al. (2011) Effect on transmission of HIV-1 resistance of timing of implementation of viral load monitoring to determine switches from first to second-line antiretroviral regimens in resource-limited settings. AIDS 25: 843–850. doi:10.1097/QAD.0b013e328344037a
67. Wagner BG, Kahn JS, Blower S (2010) Should we try to eliminate HIV epidemics by using a "Test and Treat" strategy? AIDS 24: 775–776. doi:10.1097/QAD.0b013e3283366782
68. Boily M-C, Baggaley RF, Wang L, Masse B, White RG, et al. (2009) Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect Dis 9: 118–129. doi:10.1016/S1473-3099(09)70021-0
69. Hollingsworth TD, Anderson RM, Fraser C (2008) HIV-1 transmission, by stage of infection. J Infect Dis 198: 687–693. doi:10.1086/590501
70. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, et al. (2005) Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. J Infect Dis 191: 1403–1409. doi:10.1086/429411
71. Cohen MS, Dye C, Fraser C, Miller WC, Powers KA, et al. (2012) HIV treatment as prevention: debate and commentary—will early infection compromise treatment-as-prevention strategies? PLoS Med 9: e1001232. doi:10.1371/journal.pmed.1001232
72. Powers KA, Ghani AC, Miller WC, Hoffman IF, Pettifor AE, et al. (2011) The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet 78: 256–268. doi:10.1016/S0140-6736(11)60842-8
73. Rosen S, Fox MP (2011) Retention in HIV care between testing and treatment in sub-Saharan Africa: a systematic review. PLoS Med 8: e1001056. doi:10.1371/journal.pmed.1001056
74. Cornell M, Grimsrud A, Fairall L, Fox MP, van Cutsem G, et al. (2010) Temporal changes in programme outcomes among adult patients initiating antiretroviral therapy across South Africa, 2002–2007. AIDS 24: 2263–2270. doi:10.1097/QAD.0b013e32833d45c5
75. The US President's Emergency Plan for AIDS Relief (2011 August) Guidance for the prevention of sexually transmitted HIV infections. Available: http://www.pepfar.gov/documents/organization/171303.pdf. Accessed 19 March 2012.
76. Maher D, Biraro S, Hosegood V, Isingo R, Lutalo T, et al. (2010) Translating global health research aims into action: the example of the ALPHA network. Trop Med Int Health 15: 321–328. doi:10.1111/j.1365-3156.2009.02456.x
77. Stover J, Bollinger L, Avila C (2011) Estimating the impact and cost of the WHO 2010 recommendations for antiretroviral therapy. AIDS Res Treat 2011: 738271. doi:10.1155/2011/738271
78. Stover J, Johnson P, Hallett T, Marston M, Becquet R, et al. (2010) The Spectrum projection package: improvements in estimating incidence by age and sex, mother-to-child transmission, HIV progression in children and double orphans. Sex Transm Infect 86(Suppl 2): ii16–ii21. doi:10.1136/sti.2010.044222
79. Tanser F, Hosegood V, Bärnighausen T, Herbst K, Nyirenda M, et al. (2008) Cohort profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. Int J Epidemiol 37: 956–962. doi:10.1093/ije/dym211
80. AIDS 2031 Commission (2010) AIDS in an uncertain world: the report of the AIDS 2031 Commission. Upper Saddle River (New Jersey): Prentice Hall.
81. Long E, Brandeau M (2010) The cost-effectiveness and population outcomes of expanded HIV screening and antiretroviral treatment in the United States. Ann Intern Med 158: 778–789.
82. Hontelez JA, Lurie MN, Newell ML, Bakker R, Tanser F, et al. (2011) Ageing with HIV in south africa. AIDS 25: 1665–1667. doi:10.1097/QAD.0b013e32834982ea
83. Hontelez JA, Nagelkerke N, Bärnighausen T, Bakker R, Tanser F, et al. (2011) The potential impact of RV144-like vaccines in rural South Africa: a study using the STDSIM microsimulation model. Vaccine 29: 6100–6106. doi:10.1016/j.vaccine.2011.06.059
84. Johnson L, Dorrington R, Bradshaw D, Pillay-Van Wyk V, Rehle T (2009) Sexual behaviour patterns in South Africa and their association with the spread of HIV: insights from a mathematical model. Demogr Res 21: 289–340. doi:10.4054/DemRes.2009.21.11

Editors' Summary

Background. Following the first reported case of AIDS in 1981, the number of people infected with HIV, the virus that causes AIDS, increased rapidly. In recent years, the number of people becoming newly infected has declined slightly, but the virus continues to spread at unacceptably high levels. In 2010 alone, 2.7 million people became HIV-positive. HIV, which is usually transmitted through unprotected sex, destroys CD4 lymphocytes and other immune system cells, leaving infected individuals susceptible to other infections. Early in the AIDS epidemic, half of HIV-infected people died within eleven years of infection. Then, in 1996, antiretroviral therapy (ART) became available, and, for people living in affluent countries, HIV/AIDS gradually became considered a chronic condition. But because ART was expensive, for people living in developing countries HIV/AIDS remained a fatal condition. Roll-out of ART in developing countries first started in the early 2000s. In 2006, the international community set a target of achieving universal ART coverage by 2010. Although this target has still not been reached, by the end of 2010, 6.6 million of the estimated 15 million people in need of ART in developing countries were receiving ART.

Why Was This Study Done? Several studies suggest that ART, in addition to reducing illness and death among HIV-positive people, reduces HIV transmission. Consequently, there is interest in expanding the provision of ART as a strategy for reducing the spread of HIV ("HIV treatment as prevention"), particularly in sub-Saharan Africa, where one in 20 adults is HIV-positive. It is important to understand exactly how ART might contribute to averting HIV transmission. Several mathematical models that simulate HIV infection and disease progression have been developed to investigate the impact of expanding access to ART on the incidence of HIV (the number of new infections occurring in a population over a year). But, although all these models predict that increased ART coverage will have epidemiologic (population) benefits, they vary widely in their estimates of the magnitude of these benefits. In this study, the researchers systematically compare the predictions of 12 mathematical models of the HIV epidemic in South Africa, simulating the same ART intervention programs to determine the extent to which different models agree about the impact of expanded ART.

What Did the Researchers Do and Find? The researchers invited groups who had previously developed mathematical models of the epidemiological impact of expanded access to ART in South Africa to participate in a systematic comparison exercise in which their models were used to simulate ART scale-up scenarios in which the CD4 count threshold for treatment eligibility, access to treatment, and retention on treatment were systematically varied. To exclude variation resulting from different model assumptions about the past and current ART program, it was assumed that ART is introduced into the population in the year 2012, with no treatment provision prior to this, and interventions were evaluated in comparison to an artificial counterfactual scenario in which no treatment is provided. A standard scenario based on the World Health Organization's recommended threshold for initiation of ART, although unrepresentative of current provision in South Africa, was used to compare the models. In this scenario, 80% of HIV-infected individuals received treatment, they started treatment on average a year after their CD4 count dropped below 350 cells per microliter of blood, and 85% remained on treatment after three years. The models predicted that, with a start point of 2012, the HIV incidence would be 35%–54% lower in 2020 and 32%–74% lower in 2050 compared to a counterfactual scenario where there was no ART. Estimates of the number of person-years of ART needed per infection averted (the efficiency with which ART reduced new infections) ranged from 6.3–18.7 and from 4.5–20.2 over the periods 2012–2020 and 2012–2050, respectively. Finally, estimates of the impact of ambitious interventions (for example, immediate treatment of all HIV-positive individuals) varied widely across the models.

What Do These Findings Mean? Although the mathematical models used in this study had different characteristics, all 12 predict that ART, at high levels of access and adherence, has the potential to reduce new HIV infections. However, although the models broadly agree about the short-term epidemiologic impact of treatment scale-up, their longer-term projections (including whether ART alone can eliminate HIV infection) and their estimates of the efficiency with which ART can reduce new infections vary widely. Importantly, it is possible that all these predictions will be wrong—all the models may have excluded some aspect of HIV transmission that will be found in the future to be crucial. Finally, these findings do not aim to indicate which specific ART interventions should be used to reduce the incidence of HIV. Rather, by comparing the models that are being used to investigate the feasibility of "HIV treatment as prevention," these findings should help modelers and policymakers think critically about how the assumptions underlying these models affect the models' predictions.

Additional Information. Please access these websites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed.1001245.

- This study is part of the July 2012 *PLoS Medicine* Collection, "Investigating the Impact of Treatment on New HIV Infections"
- Information is available from the US National Institute of Allergy and Infectious Diseases on HIV infection and AIDS
- NAM/aidsmap provides basic information about HIV/AIDS and summaries of recent research findings on HIV care and treatment
- Information is available from Avert, an international AIDS charity on many aspects of HIV/AIDS, including information on HIV/AIDS treatment and care, on HIV treatment as prevention, and on HIV/AIDS in South Africa (in English and Spanish)
- The World Health Organization provides information about universal access to AIDS treatment (in English, French, and Spanish); its 2010 ART guidelines can be downloaded
- The HIV Modelling Consortium aims to improve scientific support for decision-making by coordinating mathematical modeling of the HIV epidemic
- Patient stories about living with HIV/AIDS are available through Avert; the charity website Healthtalkonline also provides personal stories about living with HIV, including stories about taking anti-HIV drugs and the challenges of anti-HIV drugs

PLOS | MEDICINE

www.plosmedicine.org