Cost-Effectiveness of HIV Treatment as Prevention in Serodiscordant Couples


BACKGROUND
The cost-effectiveness of early antiretroviral therapy (ART) in persons infected with human immunodeficiency virus (HIV) in serodiscordant couples is not known. Using a computer simulation of the progression of HIV infection and data from the HIV Prevention Trials Network 052 study, we projected the cost-effectiveness of early ART for such persons.

METHODS
For HIV-infected partners in serodiscordant couples in South Africa and India, we compared the early initiation of ART with delayed ART. Five-year and lifetime outcomes included cumulative HIV transmissions, life-years, costs, and cost-effectiveness. We classified early ART as very cost-effective if its incremental cost-effectiveness ratio was less than the annual per capita gross domestic product (GDP; $8,100 in South Africa and $1,500 in India), as cost-effective if the ratio was less than three times the GDP, and as cost-saving if it resulted in a decrease in total costs and an increase in life-years, as compared with delayed ART.

RESULTS
In South Africa, early ART prevented opportunistic diseases and was cost-saving over a 5-year period; over a lifetime, it was very cost-effective ($590 per life-year saved). In India, early ART was cost-effective ($1,800 per life-year saved) over a 5-year period and very cost-effective ($530 per life-year saved) over a lifetime. In both countries, early ART prevented HIV transmission over short periods, but longer survival attenuated this effect; the main driver of life-years saved was a clinical benefit for treated patients. Early ART remained very cost-effective over a lifetime under most modeled assumptions in the two countries.

CONCLUSIONS
In South Africa, early ART was cost-saving over a 5-year period. In both South Africa and India, early ART was projected to be very cost-effective over a lifetime. With individual, public health, and economic benefits, there is a compelling case for early ART for serodiscordant couples in resource-limited settings. (Funded by the National Institute of Allergy and Infectious Diseases and others.)
I

N THE HUMAN IMMUNODEFICIENCY VIRUS (HIV) Prevention Trials Network (HPTN) 052 study,1 early antiretroviral therapy (ART), as compared with delayed ART, was associated with a 96% relative reduction in the rate of linked transmission among serodiscordant couples during 24 months of follow-up. (Early ART was defined as therapy initiated when the CD4+ T-cell count ranged from 350 to 550 per cubic millimeter, and delayed ART as therapy initiated when the CD4+ count was <250 per cubic millimeter.) The trial also showed clear clinical benefits for patients with HIV infection in the early-ART group, as compared with those in the delayed-ART group, since the early initiation of therapy prevented both primary and secondary complications of HIV infection.2 Results from observational studies have also supported the efficacy of early HIV treatment.3-6

Although the clinical efficacy of treatment as prevention may be consistent across settings, the economic value may differ, owing to differences in resource utilization, labor costs, and ability to pay (as it relates to different values of the national gross domestic product [GDP]). We projected the cost-effectiveness of early ART, as compared with delayed ART, among serodiscordant couples, accounting for both treatment and transmission effects in South Africa and India, two of the nine countries that were included in the HPTN 052 study.

METHODS

ANALYTIC OVERVIEW

In collaboration with the HPTN 052 investigators, the Cost-Effectiveness of Preventing AIDS Complications (CEPAC) International Group conducted a model-based analysis that closely mirrored the HPTN 052 trial. We used a microsimulation model of HIV disease, treatment, and transmission7-10 to compare early ART with delayed ART in serodiscordant couples. We conducted analyses to determine whether regional differences in South Africa and India had an effect on the value of early ART versus delayed ART. We evaluated the following outcomes from a modified societal perspective (excluding time and productivity costs): clinical outcomes (the rate of survival and per-person life expectancy in life-years), transmission outcomes (first- and second-order transmission events), and economic outcome (per-person HIV-related health care costs). We first simulated the trial period (24 months of follow-up) and then projected these outcomes over a 5-year period and over a lifetime (i.e., until every member of the simulated index cohort died).

We calculated the incremental cost-effectiveness ratio of early ART, as compared with delayed ART, as the change in cost divided by the change in life-years, accounting for both the health gains and medical costs for the treated patient and for the timing, decreased survival, and increased costs of the projected transmission events (see the Supplementary Appendix, available with the full text of this article at NEJM.org). We classified early ART as very cost-effective if the incremental cost-effectiveness ratio was less than the per capita GDP (three times the GDP.11-13 Early ART was considered to be cost-saving, as compared with delayed ART, when it led to both a decrease in total costs and an increase in life-years. For cost-effectiveness calculations, costs and life expectancy of both index and transmitted cases were discounted (i.e., given less value in the future) at an annual rate of 3%;12 the event outcomes (number of transmissions and percent survival) were not discounted.

CEPAC INTERNATIONAL MODEL

Disease Model

The CEPAC International model is a microsimulation of HIV disease progression and treatment that is focused on resource-limited settings.7-10 Simulated patients are generated with random draws of characteristics from distributions of age, sex, CD4+ count, and HIV RNA level (see the Supplementary Appendix). The HIV RNA level determines the monthly change in the CD4+ count. A lower count leads to a higher risk of setting-specific opportunistic diseases and HIV-related mortality. The model simulates a patient’s underlying CD4+ count and HIV RNA level; however, clinical decisions are based on presentation with an opportunistic disease or on an observed CD4+ count and HIV RNA level, which are measured on periodic laboratory testing (see the Supplementary Appendix).

According to the HPTN 052 study protocol, CD4+ counts and HIV RNA levels in the model were monitored quarterly; when a patient met strategy-specific treatment criteria (e.g., a CD4+
Cost-Effectiveness of HIV Treatment as Prevention

engl j med 369;18 nejm.org october 31, 2013

Because there were significant differences between the two countries with respect to patients’ sex and HIV RNA distributions, we used country-specific data for these variables (Table 1).14-21

Natural History
We derived monthly rates of opportunistic diseases from the widely reported Cape Town AIDS Cohort in South Africa and the Y.R. Gaitonde Centre for AIDS Research and Education (YRG CARE) cohort in India.22,23 For patients who were receiving treatment, the use of trimethoprim–-sulfamethoxazole (which was initiated at a CD4+ count of <200 per cubic millimeter)1,14,24,25 attenuated the frequency of some events.26

ART Efficacy and Loss to Follow-up
We derived data on the efficacy of ART and loss to follow-up from the trial (see the Supplementary Appendix). At 48 weeks, the rate of virologic suppression was 92%. The rate of later virologic failure was 1.4 cases per 100 person-years. The rate of loss to follow-up was 3.4 cases per 100 person-years.

Transmission
Data from the HPTN 052 study showed that the average number of secondary partners per month was 0.014 in South Africa and 0.007 in India. Within a serodiscordant partnership, trial-derived transmission rates depended on the HIV RNA level of the index patient, with rates ranging from 0.10 cases per 100 person-years for patients with virologic suppression (<400 copies per milliliter) to 1.48 per 100 person-years for patients without virologic suppression. We assumed that sexual activity was age-dependent18 and that the average duration of the primary partnership was 15 years.15 We also assumed that for the incident (first-order) cases that were not in primary partnerships, the average number of partners per month was 2.0.27 We modeled a 5-month duration of acute infection, during which infectivity increased by a factor of 26, as compared with infectivity among patients without HIV RNA suppression.16,17

Costs
We calculated the costs of treating a patient with HIV infection by multiplying resources used (e.g., inpatient days, outpatient visits, and laboratory tests) by country-specific unit costs that included costs for professional time, supplies, and non-ART medications, as well as capital costs; these values

INPUT DATA
Cohort
Characteristics of the study cohort reflected the characteristics of all participants in the HPTN 052 study. For patients in both South Africa and India, we used pooled estimates of mean (±SD) age (33.7±8.5 years) and initial CD4+ count (449±120 per cubic millimeter). Because there

Transmission Module
We used a transmission module to tally monthly transmission events to primary and secondary partners and to project both the incident transmission from the index patient (i.e., first-order transmission) and subsequent transmissions to other uninfected partners (second-order transmission) (for details, see the Supplementary Appendix). The module allowed for variation in the duration of primary partnerships and in sexual activity outside those partnerships, age-adjusted sexual behaviors, and transmissibility. For incident HIV cases, acute infection increased infectivity to future uninfected partners (see the Supplementary Appendix). On resolution of acute disease, previously uninfected incident cases — once HIV was detected — were eligible for the same strategy (early vs. delayed ART) as the index case. In the case of secondary partners, we assumed that early ART would benefit all partners of the HIV-infected person.

Downloaded from nejm.org by GE GRAY on December 4, 2013. For personal use only. No other uses without permission. Copyright © 2013 Massachusetts Medical Society. All rights reserved.
were derived from data obtained from sources other than the HPTN 052 study.\textsuperscript{19,20} We derived ART costs from the World Health Organization (WHO) Global Price Reporting Mechanism database (Table 1).\textsuperscript{21} We used country-specific GDP deflators and the average 2011 exchange rate to convert all currencies to 2011 U.S. dollars.\textsuperscript{11,28}

**Sensitivity and Scenario Analyses**

We first conducted an analysis for the base-case scenario, as defined by the input values described in Table 1. We then conducted one-way sensitivity analyses of clinical, transmission-related, and cost variables to examine their effect on four outcomes of delayed ART, as compared with early ART: the 5-year and lifetime change in cumulative transmissions and the 5-year and lifetime incremental cost-effectiveness ratio.

Since clinical trials probably represent the best-case scenario with regard to ART efficacy, transmissions, and other outcomes, we examined two additional scenarios (multiway sensitivity analyses) (Table S2 in the Supplementary Appendix). In the routine-care scenario, we used a later presentation for care (mean CD4+ count per cubic millimeter, 195 in South Africa and 305 in India).\textsuperscript{23,29} For the two countries, we used a 48-week virologic suppression rate of 74%, a later virologic failure rate of 9.5 cases per 100 person-years,\textsuperscript{30} a rate for loss to follow-up of 10.0 cases per 100 person-years through the first year of ART (and 3.2 per 100 person-years thereafter),\textsuperscript{31–33} and a slightly higher transmission rate (0.16 to 9.03 cases per 100 person-years).\textsuperscript{34} In the worst-case scenario, we incorporated cohort characteristics that were similar to those in the rou-

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base-Case Value\textsuperscript{†}</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohort characteristic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td>33.7±8.5</td>
<td>33.7±8.5</td>
</tr>
<tr>
<td>Initial CD4+ count (cells/mm\textsuperscript{3})</td>
<td>449±120</td>
<td>449±120</td>
</tr>
<tr>
<td>Sex (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>42</td>
<td>67</td>
</tr>
<tr>
<td>Female</td>
<td>58</td>
<td>33</td>
</tr>
<tr>
<td><strong>Efficacy of first- and second-line ART</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients with HIV RNA suppression at 48 wk (%)\textsuperscript{‡}</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>Increase in CD4+ count at 48 wk (cells/mm\textsuperscript{3})</td>
<td>206</td>
<td>206</td>
</tr>
<tr>
<td>Rate of failure after 48 wk (no./100 person-yr)</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Rate of loss to follow-up (no./100 person-yr)</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>Transmission</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average no. of secondary partners/mo\textsuperscript{§}</td>
<td>0.014</td>
<td>0.007</td>
</tr>
<tr>
<td>No. of transmissions/partner/100 person-yr\textsuperscript{¶}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients with suppression of HIV RNA</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Patients without suppression of HIV RNA</td>
<td>1.48</td>
<td>1.48</td>
</tr>
<tr>
<td>Average duration of primary partnership (yr)</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Average duration of acute infection (mo)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Infectivity multiplier for acute infection\textsuperscript{‖}</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Relative sexual activity according to age\textsuperscript{**}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–24 yr</td>
<td>1.10</td>
<td>1.10</td>
</tr>
<tr>
<td>25–29 yr</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>30–34 yr</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt;35 yr</td>
<td>0.82</td>
<td>0.82</td>
</tr>
</tbody>
</table>
tine-care scenario: a 48-week virologic suppression rate of 65%, a later failure rate of 14.1 cases per 100 person-years,35 a rate of loss to follow-up of 10.0 cases per 100 person-years in the first year of ART (and 5.1 per 100 person-years thereafter),31-33 and a transmission rate of 1.25 to 21.83 cases per 100 person-years.18

### RESULTS

#### SIMULATION OF TRIAL PERIOD

On the basis of clinical and ART data from the HPTN 052 study, the model produced a plausible replication of trial-based, 24-month immunologic outcomes (Fig. S1 in the Supplementary Appendix).

#### PROJECTED OUTCOMES FOR INDEX PATIENTS

##### Clinical Factors

In South Africa, the 5-year rate of survival among index patients increased with early ART, as compared with delayed ART (93% vs. 83%). Over a lifetime, the life expectancies with early ART, as compared with delayed ART, were 189.4 months and 165.2 months, respectively. Clinical outcomes in India were qualitatively similar to those in South Africa (Table 2).

### Costs

In South Africa, early ART prevented many opportunistic diseases (e.g., tuberculosis), thus preventing comparatively high costs of care (Fig. S2 in the Supplementary Appendix), a savings that more than outweighed ART costs. Consequently, over a 5-year period, early ART and delayed ART in South Africa had similar per-person costs (approximately $4,800). Over a lifetime, the longer survival that was associated with early ART increased total costs, as compared with delayed ART ($16,600 vs. $15,100) (Table 2).

In India, the average costs associated with medical care for opportunistic diseases, as compared with costs associated with ART, are lower than those in South Africa (Fig. S2 in the Supplementary Appendix). Although early ART prevented opportunistic diseases in India, the cost savings did not outweigh the ART costs; thus,
early ART led to higher total costs than delayed ART over both a 5-year period ($2,200 vs. $1,800) and a lifetime ($8,400 vs. $7,300) (Table 2).

**Transmissions**

In South Africa, early ART resulted in an immediate, large decrease in the number of projected transmissions (Fig. 1A, and Fig. S4A in the Supplementary Appendix). The difference in transmission rates was attenuated over time, as more patients started receiving therapy in the delayed-ART group. By year 6 after presentation, rates of annual transmissions were greater in the early-ART group than in the delayed-ART group because most patients in the delayed-ART group were receiving ART and some patients in the

### Table 2. HIV Transmission, Life Expectancy, Costs, and Cost-Effectiveness.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Index Patients</th>
<th>Patients with Newly Transmitted HIV</th>
<th>Overall Incremental Cost-Effectiveness Ratio†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Survival %</td>
<td>Life Expectancy mo</td>
<td>Costs 2011 U.S. $ no. (% change)</td>
</tr>
<tr>
<td>South Africa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed ART</td>
<td>83</td>
<td>51.7†</td>
<td>4,800†</td>
</tr>
<tr>
<td>Early ART</td>
<td>93</td>
<td>54.9†</td>
<td>4,800†</td>
</tr>
<tr>
<td>Lifetime period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed ART</td>
<td>NA</td>
<td>165.2</td>
<td>15,100</td>
</tr>
<tr>
<td>Early ART</td>
<td>NA</td>
<td>189.4</td>
<td>16,600</td>
</tr>
<tr>
<td>Alternative scenarios</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Routine care</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst case</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed ART</td>
<td>87</td>
<td>52.8†</td>
<td>1,800</td>
</tr>
<tr>
<td>Early ART</td>
<td>94</td>
<td>55.1†</td>
<td>2,200</td>
</tr>
<tr>
<td>Lifetime period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed ART</td>
<td>NA</td>
<td>182.6</td>
<td>7,300</td>
</tr>
<tr>
<td>Early ART</td>
<td>NA</td>
<td>202.6</td>
<td>8,400</td>
</tr>
<tr>
<td>Alternative scenarios</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Routine care</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worst case</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* NA denotes not applicable.
† The overall incremental cost-effectiveness ratio was calculated as the difference in costs for the index patient plus the difference in the increase in costs due to transmission divided by the difference in life expectancy for the index patients plus the difference in life-months lost due to transmission, all discounted at 3% per year.
‡ The percent change in transmissions per index patient is indicated for early ART, as compared with delayed ART, and was calculated before rounding.
§ Life expectancy for the 5-year period indicates the average duration of survival through 5 years.
¶ Costs are the same for delayed and early ART because of rounding. Unrounded values are $4,766 for delayed ART and $4,780 for early ART.
‖ Additional details of the results of the routine-care and worst-case scenarios are provided in Tables S3 and S4, respectively, in the Supplementary Appendix.
early-ART group had begun to have virologic failure. However, rates of cumulative transmissions in the early-ART group remained consistently lower than those in the delayed-ART group throughout the lifetime of the index patient (Fig. 1B). Projected transmission results were similar in the two countries (Fig. S3A, S3B, and S4B in the Supplementary Appendix).

Cost-Effectiveness
When we included transmission-related life-months lost and costs accrued over a 5-year period, early ART was cost-saving, as compared with delayed ART, in South Africa. Over a lifetime, early ART was very cost-effective ($590 per life-year saved) (Table 2, and Fig. S6 in the Supplementary Appendix). In India, early ART was cost-effective over a 5-year period ($1,800 per life-year saved) and very cost-effective over a lifetime ($530 per life-year saved) (Table 2, and Fig. S6 in the Supplementary Appendix).

In the two countries and over the two time periods, the clinical benefit to the index patient who received early ART was much larger in magnitude than the clinical benefit associated with the prevention of new transmissions. For example, for a patient receiving early ART in South Africa, the 5-year increase in life expectancy of 3.2 months (from 51.7 to 54.9 months) was much larger than the clinical benefit associated with the prevention of new transmissions, which resulted in 0.07 life-months gained (a decrease in life-months lost from 0.09 to 0.02).

SENSITIVITY ANALYSES

Transmissions
In the South African base-case scenario, early ART, as compared with delayed ART, resulted in a relative reduction of 69% in the number of transmissions over a 5-year period and of 13% over a lifetime (Table 2). Input values related to treatment — specifically, changes in 48-week suppression rates, late virologic failure (after initial suppression), and loss to follow-up — had the greatest effect on these results (Fig. S7A and S7B in the Supplementary Appendix). Regardless of these variations, early ART consistently resulted in fewer infections than delayed ART over the 5-year and lifetime periods. Similar results were seen in India (Fig. S7C and S7D in the Supplementary Appendix).

Cost-Effectiveness
In South Africa, in most of the ranges that we examined, early ART was cost-saving over a 5-year period. Variation in several treatment-related variables (including increased rates of late failure associated with poor ART adherence) moved the incremental cost-effectiveness ratio for early ART from cost-saving to very cost-effective; however, for transmission-related variables, these ratios were consistently cost-saving (Fig. 2A). Five-year results were also sensitive to these variables in India, where some variation in treatment-related variables moved
Figure 2. Results of One-Way Sensitivity Analyses.

Shown are the effects of various treatment and transmission variables on the incremental cost-effectiveness ratio (the change in cost divided by the change in life-years) for early ART, as compared with delayed ART, over a 5-year period in South Africa (Panel A) and India (Panel B). ART efficacy was defined as HIV RNA suppression at 48 weeks. In these tornado diagrams (so named because the data are listed vertically in a funnel pattern, with the most important variables at the top), the solid vertical line indicates the incremental cost-effectiveness ratio for the base-case scenario in the study. The length of the horizontal bars indicates the range of incremental cost-effectiveness ratios over the specified range of variable values; the longer the bar, the greater the sensitivity of the incremental cost-effectiveness ratio to that variable. To the right of the horizontal bars in Panel A, the range of values that were examined is shown in parentheses, with the value producing the lower incremental cost-effectiveness ratio listed first. Incremental cost-effectiveness ratios that met the definition of cost-saving (which was defined as both a decrease in total costs and an increase in life-years) are not shown in the tornado diagram. Instead, the cost-saving range for incremental cost-effectiveness ratios is provided to the left of the horizontal bars. The dashed vertical black line indicates the gross domestic product (GDP) of the country (the threshold for being very cost-effective), and the dashed vertical red line represents three times the GDP of the country (the threshold for being cost-effective, which is shown for India only). The GDPs are shown in dollars (see the Methods section). In South Africa, numerous ranges of values that were examined, including all transmission variables, showed that early ART was cost-saving over a 5-year period. In India, early ART, as compared with delayed ART, was found to be cost-effective over a 5-year period for the base case; with variation in some treatment variables, early ART was very cost-effective (<$1,500 per life-year saved). OD denotes opportunistic disease, and PY person-year.

**SCENARIO ANALYSES**

In the two countries and for both time periods, projected clinical outcomes were worse for alternative scenarios in which there were lower rates of efficacy for ART, higher rates of late virologic failure, and higher rates of loss to follow-up (see Table S3 for the routine-care scenario and Table S4 for opportunistic disease, and PY person-year.

**early ART from cost-effective to very cost-effective (Fig. 2B).** In one-way sensitivity analyses, early ART was very cost-effective over a lifetime across all ranges of variables that were examined in the two countries (Fig. S8A and S8B in the Supplementary Appendix).
for the worst-case scenario in the Supplementary Appendix). With later ART initiation, less effective treatment, and higher transmission levels, early ART had more modest benefits with respect to transmission at 5 years, with no benefit over a lifetime (Fig. S5 in the Supplementary Appendix). Despite these changes in variables and subtle differences in outcome, economic and policy conclusions were robust. In the base-case scenario, as well as in the routine-care and worst-case scenarios, early ART at 5 years was cost-saving in South Africa and cost-effective in India. In all scenarios and in both countries, early ART was very cost-effective over a lifetime (Tables S3 and S4 in the Supplementary Appendix).

**DISCUSSION**

With compelling evidence supporting the clinical and preventive benefits of early ART, the focus is now on economic questions with respect to early ART.3–6 In a model-based analysis projecting the clinical, population, and economic effects of the HPTN 052 study in South Africa and India, we found that early ART substantially improved the rate of survival of infected patients, greatly decreased the rate of early HIV transmissions, and provided an excellent return on investment. We specifically chose to model these two countries because they had the highest numbers of persons with HIV infection among the nine countries that were evaluated in the HPTN 052 study and because South Africa and India were representative of middle-income and lower-income countries, respectively (Table S5 in the Supplementary Appendix).11 Over the first 5 years, the differences in costs of care and GDP accounted for differences in the cost-effectiveness of early ART. In South Africa, early ART was found to be cost-saving, whereas in India, early ART was cost-effective. Over the long term, early ART was very cost-effective in both countries.

Data from the HPTN 052 study highlight the benefits of early ART with respect to both morbidity and transmission. Early ART significantly increased the time to a diagnosis of tuberculosis, a WHO stage 4 event, or secondary complications, as compared with delayed ART.2 In our model-based, longer-term projections, these delays in morbidity led to substantial reductions in mortality — differences that were not detected in the short timeframe of the HPTN 052 study. Indeed, these data and projections underscore that among the greatest prevention effects of early ART is the durable prevention of opportunistic diseases.

Our analyses show that the preventive value of early ART depends on the extent to which patients remain in care and receive effective treatment. The capacity to maintain virologic suppression among patients who are retained in care will determine the magnitude of long-term benefits with respect to transmission.

Although our study shows that early ART is very cost-effective, a major challenge to its implementation — in addition to the challenge of earlier HIV disease detection36 — is affordability in resource-limited settings. This should not be the case in South Africa. Over the first 5 years, the cost of early ART is more than offset by the prevention of common, expensive-to-treat opportunistic diseases; the opportunity to offset those costs is lost when ART is delayed. In India, upfront expenditures will be needed to implement early ART.

Our results are more conservative — with respect to cumulative transmissions prevented by early ART — than those modeled in studies reported previously.37,38 The reason for this is that we accounted for the individual benefit of early ART and the monthly changes in virologic-based transmissions associated with HIV detection, virologic suppression and rebound, and loss to follow-up. We also specifically chose to examine countries with different care and cost structures in order to show the generalizability of these results in two very different resource-limited settings.

In the HPTN 052 study, a total of 39 HIV transmission events were reported, with definitive HIV phylogenetic data available for 36 of these events. Of the latter transmission events, 7 (19%) were not phylogenetically linked to the primary partner and were presumed to have been acquired from secondary partners,39 whose characteristics are unknown. If many of these secondary partners would not be eligible for ART under a policy of early ART for serodiscordant couples, then we may have overestimated the preventive effect of this policy. Alternatively, if the secondary partners would be eligible for ART as members of serodiscordant couples,
then we may have underestimated the preventive effect of early ART. Regardless of the direction of this bias, our results and sensitivity analyses showed that early ART would be very cost-effective.

Our study has several limitations. First, we intentionally adhered closely to the design, population, and data of the HPTN 052 study. Although we have examined the potential generalizability of the conclusions in routine-care and worst-case scenarios, we cannot be sure that the results are representative of care in nontrial settings or of patients with HIV who are not in regular serodiscordant partnerships. Second, in the scenario analyses, we did not consider the possibility of increased drug resistance because of poor adherence. Third, since HIV-infected patients were already identified in the HPTN 052 study, we consider our study to be an analysis of “treatment as prevention” rather than an analysis of “test and treat.” Fourth, although we excluded productivity and other nonmedical economic benefits of prevented transmissions, the inclusion of such factors would make the results even more economically attractive. Fifth, despite the many benefits of modeling HIV progression together with transmissions, such a complex model is not easily conducive to probabilistic sensitivity analysis. Finally, we limited the analysis to first-order and second-order transmissions from the index patient; as indicated by the modest effect of second-order transmissions, the inclusion of more distal transmissions would no doubt have a minimal effect on the results.

In conclusion, our analysis using the clinical data from the HPTN 052 study and published data from cohorts in South Africa and India shows that early ART provides substantial individual and population benefits over delayed ART. The ultimate preventive value of early ART is a shared responsibility of providers and patients, and patients who remain in care with virologic suppression will glean the maximum long-term preventive effect of early ART. Early ART also has major economic benefits and is very cost-effective — results that are robust across settings and assumptions.

Supported by grants from the National Institute of Allergy and Infectious Diseases (NIAID) (R01 AI058736, R01 AI093269, and UM1 AI068636) and the National Institute of Mental Health (R01 MH087328) (all to the CEPAC International Group), by the HIV Prevention Trials Network (HPTN), and by grants from NIAID (UM1 AI068619, U01 AI068619, UM1 AI068613, and U01 AI068631, to the HPTN Network Laboratory; and U01 AI068617 and U01 AI068617, to the HPTN Statistical and Data Management Center) and by a grant from the AIDS Clinical Trials Group (UM1 AI068636). Study drugs for the HPTN 052 trial were donated by Abbott Laboratories, Boehringer Ingelheim, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline/III/Healthcare, and Merck.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the HPTN 052 study team and the study participants.

APPENDIX

The authors’ affiliations are as follows: the Divisions of Infectious Disease (R.P.W., K.A.F.) and General Medicine (R.P.W., E.L.R., F.N., Y.M.N., E.L., K.A.F.), the Medical Practice Evaluation Center (R.P.W., E.L.R., F.N., Y.M.N., E.L., K.A.F.), Department of Medicine, Massachusetts General Hospital, the Division of Infectious Disease (R.P.W.) and the Department of Orthopaedic Surgery (E.L.), Brigham and Women’s Hospital, the Harvard University Center for AIDS Research, Harvard Medical School (R.P.W., E.L., K.H.M., K.A.F.), the Departments of Health Policy and Management (M.C.W., K.A.F.) and Epidemiology (G.R.S.), and the Center for Biostatistics in AIDS Research (G.R.S.), Harvard School of Public Health, and Fenway Health (K.H.M.) — all in Boston; the Y.R. Gaitonde Center for AIDS Research and Education, Chennai (N.K.), and National AIDS Research Institute, Pune (S.V.G.) — both in India; the Desmond Tutu HIV Centre, Institute of Infectious Disease and Molecular Medicine, and Department of Medicine, University of Cape Town, Cape Town, South Africa (R.W.); the Yale School of Public Health, New Haven, CT (A.D.P.); Perinatal HIV Research Unit (R.P.) and Clinical HIV Research Unit (I.S.), Department of Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg; the Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle (T.G., L.W.); FHI 360, Washington, DC (M.M.); FHI 360, Durham, NC (T.G.); and University of North Carolina School of Medicine, Chapel Hill (M.S.C.).

REFERENCES

8. Goldie SJ, Yazdanpanah Y, Losina E,


Copyright © 2013 Massachusetts Medical Society.