INTRODUCTION

The beneficial effects of antiretroviral therapy (ART) on individual health are well established, and ART is widely used to reduce the morbidity and mortality due to the human immunodeficiency virus (HIV) (WHO 2016). Recent evidence has strengthened the case for initiating ART as early in the disease stage as possible (Danel and others 2015; INSIGHT START Study Group 2015). Similarly, using ART to prevent mother-to-child transmission of HIV is supported with conclusive evidence and has been adopted into clinical policies worldwide, as discussed in chapter 6 of this volume (John-Stewart and others 2017). Years of accumulating biological and observational evidence also suggest that ART may reduce sexual transmission of HIV, although the field lacked conclusive evidence until recently (Donnell and others 2010; Nachega and others 2013).

The evidence base and attention to “treatment as for prevention” strengthened substantially in 2011 with the interim results from HIV Prevention Trials Network (HPTN) 052, a randomized controlled trial of early versus delayed use of ART among serodiscordant couples (Cohen and others 2011). The trial demonstrated a 96 percent reduction in new infections with earlier initiation of ART and provided strong evidence that ART reduces the sexual transmission of HIV. Final results of this trial with nearly 10,000 person-years of follow-up with similar conclusions were published in 2016 (Cohen and others 2016).

This emerging evidence stimulated a range of questions regarding the biological mechanisms of HIV treatment as prevention (TasP), variations in efficacy across subgroups, differences in at-risk populations, optimal implementation strategies, and potential implications for public health (Cohen, Holmes, and others 2012; Delva and others 2012). Recognition of the dual benefits of treatment has resulted in the reevaluation of the cost-effectiveness of ART, as well as of the paradigms of HIV prevention (Garnett and others 2017) and has led to policy discussions about how best to value the risks and benefits of treatment for personal and public health.

Even as substantial research and evaluation have improved the understanding of these trade-offs, clinical and public health policy and funding decisions are being made at the program, national, and global levels. This chapter examines the concept of HIV TasP, focusing on the underlying biological mechanisms, effectiveness, and cost-effectiveness of various strategies and settings and assessing how these factors may influence resource allocation, policy decisions, and research agendas at the national and global levels.
THE BIOLOGY OF TRANSMISSION

HIV is transmitted in three ways: from parenteral exposure to contaminated blood and blood products; from exposure of many mucosal sites to infected genital secretions; and from mother to baby before, during, or after delivery (Royce and others 1997). These routes of transmission have been studied extensively and found to have different probabilities of transmission, given exposure. In each case, the biology of transmission is believed to be defined by the infectiousness of the host and the susceptibility of the person exposed (Cohen and others 2010; Cope and others 2014; Pilcher and others 2004).

HIV-infected fluids contain cells infected and not infected with HIV. The replication of HIV generates a very large number of viral variants—viruses that have different nucleic acid sequences—that constitute an infectious swarm. Within the swarm, some viruses are capable of producing infection; others are defective or less fit for transmission (Ho and others 2013). The likelihood that HIV will cause infection is governed by the number of viruses—the inoculum (Baeten and others 2011; Donnell and others 2010; Laeyendecker and others 2012; Quinn and others 2000)—and the genotypic and phenotypic characteristics of HIV in the swarm (Martin and others 2014). The probability of HIV transmission in heterosexual couples directly reflects the concentration of HIV in the fluid studied (Baeten and others 2011; Quinn and others 2000). In a landmark study of heterosexual transmission, Quinn and others (2000) observed no transmission when the blood plasma viral load was less than 1,500 copies per milliliter, and the most transmission when the viral load was more than 37,500 copies per milliliter. Unprotected anal intercourse appears to have high risk of transmission per contact (Baggaley, Dimitrov, and others 2013), explaining the high incidence of HIV in men who have sex with men (MSM), but the viral load required for transmission by this route has not been determined.

The phenotype of the founder viruses that initiate infection helps determine the probability of HIV transmission above and beyond the inoculum effects (Carlson and others 2014; Parrish and others 2013; Ping and others 2013). HIV variants that cause infection are dual tropic—that is, they use both CD4 and CCR5 receptors (Joseph and others 2014; Ping and others 2013; Shaw and Hunter 2012). Only one to three founder viruses are generally transmitted, and the number of variants may reflect the route of exposure (Keele and others 2008). Transmission from penile-vaginal exposure has the fewest variants, followed by anal exposure, followed by parenteral exposure (Li and others 2010).

Susceptibility to infection varies greatly. The only proven relative immunity to HIV results from deletion in the CCR5 receptor, which has been observed in about 1 of 100 Caucasians and less commonly in non-Caucasians (O’Brien and others 2000). Many studies have tried to define factors that allow some people to remain exposed and uninfected (McLaren and others 2013). By definition, all HIV-discordant couples include a partner who is exposed and uninfected (Muessig and Cohen 2014). Yet, many people in this group will become infected.

There is little evidence to suggest that innate immunity, antibodies, or T-cell responses provide durable or reliable resistance to HIV. More likely, “apparent resistance” reflects the absence of factors that amplify transmission (Pilcher and others 2007). Inflammation from any source will cause defects in the mucosa, evoke a large number of receptive cells, increase the number of receptors expressed, and activate cells that favor HIV replication. For example, bacterial vaginosis characterized by a change in vaginal flora and watery discharge is strongly associated with HIV acquisition in women (Taha and others 1998). Unique cytokine profiles may also favor these conditions (Olivier and others 2014).

TasP uses ART to reduce the replication of HIV in the blood and mucosal secretions profoundly, quickly, and reliably. The hypothesis that treatment could serve as prevention began as soon as the first ART was developed (Henry 1988). Numerous groups have since demonstrated the ability of ART to penetrate and suppress viral replication in the male and female genital tract (Thompson, Cohen, and Kashuba 2013). Many antiretroviral agents achieve similar or higher concentrations in the genital tract as in blood (Kwara and others 2008). However, the body’s ability to metabolize and eliminate the medications may compromise the prevention benefits of treatment, and viruses isolated from areas of low or variable drug levels have demonstrated site-specific resistance.

To date, few cases of HIV transmission have been documented when a person with HIV has been treated sufficiently to prevent viral replication (Cohen and others 2016; Rodger and others 2016). However, HIV can be found in the male and female genital tract secretions even when HIV is suppressed in the blood (Anderson and Cu-Uvin 2011; Reichelderfer and others 2000). The implication is that the HIV detected in the genital tract under these conditions may be defective and incapable of causing infection (Zhang and others 1998).

An additional concern is the pharmacology of ART in the genital tract, which is considered a special compartment. Studies of the female genital tract
and colorectum have noted that the concentrations of tenofovir and emtricitabine and their active metabo-
lites vary according to the type of mucosal tissue (Patterson and others 2011). Differential penetration
or metabolism of ART offers further insight into the highly variable level of protection conferred by these
agents (Hendrix and others 2008). The results from preexposure prophylaxis (PrEP) clinical trials suggest
that the use of ART for prevention can be optimized by choosing agents that (1) preferentially penetrate sites
of HIV acquisition or transmission or (2) have a long tissue half-life that might provide a pharmacologic
buffer for imperfect drug adherence. In summary, extensive studies of HIV transmission have been com-
pleted, and the results help illuminate the understanding of the ways to use ART to maximize the prevention
of transmission.

EVIDENCE OF THE EFFECTIVENESS OF TREATMENT AS PREVENTION

Four lines of complementary evidence support the idea that treatment in HIV-infected individuals reduces their
transmission of HIV to others:

• Observational studies of serodiscordant couples
• A randomized controlled trial
• Ecologic studies
• Population-based studies.

### Observational Studies of Serodiscordant Couples

As shown in annex 5A, 14 observational studies of sero-
discordant couples have been reported (Muessig and
Cohen 2014). In 11 of these, ART was associated with the
prevention of HIV transmission. The two studies from
China failed to note a prevention benefit from ART (Lu
and others 2010; Wang and others 2010). A larger retro-
spective analysis of 38,862 serodiscordant heterosexual
couples across China noted a 26 percent relative reduc-
tion in transmission when the index case received ART
(Jia and others 2012). In most of these studies (including
studies in China), it is not known either whether the
HIV-infected person receiving ART was actually using
the agents prescribed or what degree of viral suppression
was achieved.

Several systematic reviews of TasP studies have been
directed. Attia and others (2009) reviewed 11 cohorts
reporting on 5,021 heterosexual couples and 461 trans-
mision events. The transmission rate overall from
patients on ART was 0.46 per 100 person-years, based on
five events. The transmission rate from a seropositive
partner with a viral load less than 400 copies per milli-
fier was zero for persons on ART and 0.16 per 100 per-
son-years for persons not on ART, based on five studies
and one event. A meta-analysis of studies of serodiscor-
dant heterosexual couples where the HIV-positive part-
ner was on ART and virally suppressed found zero
transmissions per 100 person-years (Loutfy and others
2013); a similar review of partners on combination ART
for at least six months found a transmission risk of
between 1 and 13 per 100,000 sex acts. Another
meta-analysis of 50 publications found a 91 percent
(79 percent to 96 percent) reduction in incidence of
HIV-1 per partner among couples when the index case
used ART (Baggaley, White, and others 2013). Supervie
and others (2014) reported at most one HIV transmis-
sion over an estimated 113,480 sex acts—of which
17 percent were not condom protected—among 1,672
serodiscordant couples where the index partner had
been treated for more than six months.

The PARTNER Study is assessing the occurrence of
linked transmission among serodiscordant heterosexual
and MSM couples who have condomless sex, are not tak-
ing PrEP, and have a recent viral load of less than 200 cells
per cubic millimeter (cells/mm³). No linked transmis-
sions among MSM couples were observed among
1,238 couple-years of follow-up (Rodger and others
2016), implying that ART treatment prevented transmis-
sion of HIV during unprotected anal intercourse.

### The HPTN 052 Randomized Controlled Trial

HPTN 052 was a randomized controlled trial designed
to provide an understanding of the magnitude and dura-
bility of ART for prevention. The study enrolled 1,562
serodiscordant couples at 13 sites in nine countries in
Africa, Asia, and the Americas; it randomized infected
men and women to start ART at CD4 T-cell counts of
200–250 cells/mm³, compared with subjects who started
ART at CD4 T-cell counts of 350–550 cells/mm³ (median
cell count of 446 cells/mm³). All participants were
offered couples counseling for prevention. In those
receiving delayed ART, the counseling itself appeared to
reduce HIV transmission to levels far lower than in earlier
studies (less than 2 percent per year). However, the
addition of early ART led to a 96 percent prevention of
HIV transmission compared with delayed ART in an
interim analysis. Infected subjects who were treated ear-
lier not only had CD4 T-cell counts that rose quickly but
also developed fewer infections (Grinsztejn and others
2014). After 8,494 person-years of follow-up, early ART
maintained 93 percent effectiveness in the prevention of
new linked infections compared with delayed ART (Cohen and others 2015; Cohen and others 2016).
Ecological and Population-Based Studies

As shown in annex 5B, a large number of ecological studies demonstrate the ability of ART to reduce the incidence of HIV (Smith and others 2012). Most of these are from North America (Castel and others 2012; Das and others 2010; Katz and others 2002; Montaner and others 2010; Porco and others 2004; Wood and others 2009); one is from Taiwan, China (Fang and others 2004); and one is from Australia (Wand and others 2011). Each study used an ecological measure of exposure (access to ART), outcome (HIV incidence), or both. The reliability of the results lies in the strength of the measurements used for exposure and outcome.

The exposure of the entire HIV-infected population to ART can only be measured if every person infected with HIV can be identified and their treatment and virological suppression status assessed. Indeed, the hypothesis that use of ART by the entire population infected will decrease HIV incidence assumes that ongoing care will sustain viral suppression, thereby preventing transmission (Cohen and others 2011; Walensky and others 2010). In the first population-based ART randomized controlled trial completed (AAAS 2016; Iwuji, Orne-Gliemann, Larmarange, and others 2016), individuals living in communities in KwaZulu-Natal, South Africa, receiving “immediate ART” irrespective of CD4 T-cell count did not have lower incidence of HIV than those in control (standard of care) communities; however, individuals in the immediate ART communities did not have the anticipated uptake and benefits of ART because of the difficulty of implementing this strategy. Several other community randomized TasP trials are underway (Boily and others 2012).

Challenges in the Measurement of Population-Level HIV Incidence

The ability to detect a benefit of TasP depends on the ability to detect changes in HIV incidence. Widely different methods have been developed to measure HIV incidence. Perhaps the most commonly used approach estimates population-based incidence using information on newly identified cases as a proxy for new infections (Castel and others 2012; Das and others 2010; Montaner and others 2010). Newly diagnosed patients acquired HIV at some unknown earlier time, and they are not “incident” in the traditional sense. Using new diagnoses as a proxy for incidence also misses people who do not seek testing; these people may have less access to health care and a greater risk of acquiring HIV (Lopez-Quintero, Shtarkshall, and Neumark 2005; Spielberg and others 2003). Another approach is to use back-calculation from new diagnoses (Fang and others 2004), although this approach relies on assumptions related to disease progression markers such as the onset of symptoms and decline in CD4 T-cell count to estimate the time of infection (Holmes and others 2006; Novitsky and others 2010; Wand and others 2009; Wolbers and others 2010).

Longitudinal cohort follow-up data have also been used to define population incidence and are considered the gold standard of HIV incidence estimation, despite well-known sources of bias (Porco and others 2004; Wood and others 2009). In a striking example of the power of cohort studies, Tanser and others (2013) enrolled 16,000 HIV-negative people from 2005 to 2011 to receive HIV antibody testing every six months. An HIV-negative individual living in a community with 30 percent to 40 percent ART coverage was 38 percent less likely to acquire HIV than a person living in a community with less than 10 percent ART coverage. As noted above, no change in cohort incidence was observed in a cluster randomized controlled trial in the same area (Iwuji, Orne-Gliemann, Balestre, and others 2016). The likely reason for the failure of the trial to show effectiveness was that ART coverage was nearly the same in both the intervention and the control arms of the trial. Links to care were low, and the TasP intervention generally did not induce more people in the intervention arm to take up ART. In contrast, in the population-based cohort study by Tanser and others (2013), ART coverage across different geographic communities ranged from less than 10 percent to 30 percent to 40 percent.

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. The serologic testing algorithm for recent HIV seroconversion derives HIV incidence based on differences in antibodies generated in the weeks after infection (Janssen and others 1998), although logistical challenges in storing and tracking remnant blood can affect the completeness of data (Das and others 2010; Katz and others 2002). Even relatively new laboratory methods misclassify established and early infections (Le Vu and others 2008), but other methods are in development (Burns and others 2014). Currently, surveillance for recent infections in low- and middle-income countries (LMICs) is limited, which further constrains the ability to track the effect of intervention scale-up on the incidence of recent infections. However, successful
development of serological detection of incidence infection would allow cross-sectional detection of incident HIV infection either in stored samples or in demographic surveys.

**Modeling Population-Level Prevention Effectiveness**

Mathematical modeling has been used extensively to gauge the potential of ART to reduce or eliminate the spread of HIV, and virtually all models report a benefit from ART; the magnitude of the benefit reflects the degree of coverage, model assumptions, and program quality issues such as retention and adherence (Cohen and others 2013; Maddali and others 2015). In a powerful and controversial analysis of the South African HIV epidemic, Granich and others (2009) projected that massive expansion of testing and treatment (“test and treat”) along with best case program quality could substantially reduce and potentially eliminate the HIV epidemic in South Africa within 10 years. Wagner and Blower (2012) also demonstrated the theoretical possibility of HIV epidemic elimination in South Africa using a test-and-treat approach. However, they reported that it would take 40 years, and the cumulative costs would be much higher. The differences were partly attributable to differing model assumptions about survival time on ART and the costs of ART over time. A modeled analysis of expanded testing and treatment regardless of CD4 T-cell count in Washington, DC, found a more modest impact on HIV transmission (Walensky and others 2010). In a comparison of 12 independent mathematical models, Eaton and others (2012) reported broad agreement regarding the substantial potential to reduce HIV incidence in generalized epidemics in Sub-Saharan Africa, despite large differences in the structures of the models. For example, in South Africa and Zambia, expanding ART eligibility to all HIV-positive adults was projected to avert 9 percent to 40 percent of new infections over a 20-year time horizon, with greater reductions attributed to strategies involving increased testing and links to care.

Multiple investigators have also modeled the effects of various strategies incorporating TasP in concentrated epidemics in which the HIV epidemic has the largest burden among specific populations, such as persons who inject drugs (PWIDs), MSM, and female sex workers (FSWs) (Boily and Shubber 2014). In an analysis of the epidemic in Belgaum, India, the expansion of eligibility to all FSWs resulted in a 13 percent decline in projected HIV infections (Eaton and others 2012). Expanding eligibility to all HIV-positive adults, in conjunction with prioritized access for FSWs, resulted in 29 percent to 41 percent of new HIV infections being averted. In Vietnam, expanding eligibility to targeted groups produced small declines in HIV incidence: 2 percent in FSWs, 5 percent in MSM, and 5 percent in PWIDs; in contrast, expanding eligibility to all adults and prioritizing access for all three key populations resulted in a 30 percent cumulative decline in new infections (Eaton and others 2012).

A model assessing the impact of a test-and-treat strategy for urban MSM in New York City estimated a reduction in new cases of 39.3 percent over 20 years. The annual testing component of this approach provided the majority of the projected impact, whereas earlier treatment (at CD4 T-cell counts of less than 500 cells/mm³) itself contributed to an 8.5 percent decline in new infections over 20 years (Sorensen and others 2012). A test-and-treat strategy for adults with HIV in British Columbia using a model specifically built to include the main drivers of the local epidemic demonstrated 37 percent to 62 percent reductions in new infections over 25 years (Lima and others 2008).

The effectiveness of TasP will be highly dependent on the elements of the HIV cascade, as outlined in detail in chapter 4 of this volume (Harrispersaud and others 2017), including testing frequency and coverage, links to care, adherence to treatment, virological suppression, and long-term retention in care (Delva and others 2012). Maddali and others (2015) projected that moving to early treatment in India resulted in a reduction from an estimated 1,285,000 new HIV infections to 1,050,000 infections under existing program conditions over a 20-year period. However, with enhanced testing, links to care, and retention in care, the projected number of new infections with early treatment was projected to fall further to 517,000. As pointed out by Wilson and Fraser (2014), country-level data on virological suppression are variable. For example, 26.1 percent of 266 individuals reporting ART use in the 2012 Kenya AIDS Indicator Survey were found to have a detectable viral load greater than 550 copies (Cherutich and others 2016; National AIDS and STI Control Programme 2013).

Efforts to project the impact of HIV TasP strategies have highlighted the importance of understanding the relative infectiousness of people with acute and early infection. HIV transmission is more efficient during acute infection, reflecting higher viral loads and phenotypic factors that favor transmission (Cohen, Dye, and others 2012). A study in Uganda reported that people with acute and early infection are 26 times more likely to transmit HIV than people with established infection (Hollingsworth, Anderson, and Fraser 2008). Viral phylogenetic results suggest that acute and early infections are responsible for one-third to one-half of new HIV cases in MSM (Brenner, Wainberg, and Roger 2013; Rieder and others 2010). A modeling study by Eaton
and Hallett (2014) reported that a higher proportion of early infection lessened the impact of ART on estimated incidence in the first year in South Africa but did not have an important influence on the long-term effect on incidence (figure 5.1). Powers, Kretzschmar, and Miller (2014) have challenged the conclusions of this report, and the contribution of people with acute and early infection to the spread of HIV continues to be debated. The debate turns on numerous assumptions, including levels and distribution of risk behavior and epidemic patterns.

Regardless of acute infection’s potential impact, diagnosing and linking people to care as early as possible are crucial. However, it is difficult to detect and diagnose people with acute and early infection. When acutely infected patients are identified, U.S. guidelines recommend immediate treatment to preserve CD4 T-cell count, shrink the viral reservoir, and reduce HIV transmission (DHHS 2014). The World Health Organization (WHO) has not yet issued specific guidelines related to acute infection, although it does recommend treatment for all HIV-infected individuals (WHO 2016).

In summary, the effectiveness of TasP has been well established through observational studies and clinical trials. Ecological studies and projection models further demonstrate the substantial potential for population-level HIV prevention from expanded treatment across a wide variety of geographies, epidemic types, and populations. Ongoing population-based studies will further evaluate the validity of these models and provide additional evidence on the impact of TasP strategies and the real-world effects of variable program quality along the HIV treatment cascade.

**EVALUATING COST-EFFECTIVENESS**

**Metrics of Cost-Effectiveness**

Given the effectiveness of treatment for reducing the sexual transmission of HIV, it is increasingly important for policy makers to consider the cost-effectiveness of TasP. Accordingly, analysts have begun to grapple with how best to represent the range of effects of ART.

In its simplest form, a narrow definition of cost-effectiveness has been represented as the incremental cost per infection averted by ART (Bärnighausen, Salomon, and Sangruej 2012; Ying and others 2015). However, this outcome alone does not value ART’s long-term health and health-related quality-of-life effects in the denominator of the cost-effectiveness ratio as recommended by consensus guidelines for cost-effectiveness (Gold and others 1996; Weinstein and others 1996). A trial-based cost-effectiveness analysis of HPTN 052 and other analyses used joint measures—for example, life-years saved, disability-adjusted life years (DALYs) averted—to value the impacts on both health and prevention (Eaton and others 2014; Walensky and others 2013). In this construct, life-years saved or DALYs averted by ART include both direct effects on health and downstream (discounted) effects on the prevention of new infections. Other analysts have reported the cost per death averted, which similarly values deaths directly averted by the therapeutic and preventive effects of ART, although this approach is less common and does not fully account for health-related quality of life (Bärnighausen, Bloom, and Humair 2012).

**Cost-Effectiveness Estimates of Treatment**

First-generation studies evaluated the cost-effectiveness of ART versus no ART and generally did not include the prevention effects of treatment. Second-generation cost-effectiveness studies of ART examine circumstances in which ART is widely used for its health benefits. Instead of comparisons with no ART, these studies
look at expanding ART to various groups and include the effect of ART on sexual transmission (table 5.1). Most of these studies model the cost-effectiveness of earlier initiation of treatment compared with later initiation and typically include scenarios in which treatment is started at CD4 T-cell counts greater than 500 cells/mm$^3$ or is started promptly regardless of CD4 T-cell count; in contrast, the 2010 WHO guidelines (WHO 2010) recommend beginning treatment at CD4 T-cell count of 350 cells/mm$^3$. Some analyses also include related interventions, such as expanded testing and links to care, and compare ART with other prevention modalities, such as PrEP and voluntary medical male circumcision (VMMC).

These second-generation studies generally demonstrate more favorable cost-effectiveness than previous analyses. In a cost-effectiveness analysis that considered the sexual prevention–related effects of ART, Long, Brandeau, and Owens (2010) estimated that expanded treatment using prevailing eligibility criteria of CD4 T-cell levels < 200 cells/mm$^3$ was very cost-effective in the United States and that increased frequency of HIV testing resulted in a substantial additional decrease in incidence and remained very cost-effective.

In generalized epidemics in Sub-Saharan Africa, all analyses demonstrated the cost-effectiveness of further ART expansion, including early ART. Five of the 10 published analyses focused exclusively on South Africa. In an extensive analysis, Eaton and others (2014) assessed the cost-effectiveness of earlier treatment using six independent models for South Africa and four for Zambia. The models incorporated a common costing framework and some common assumptions, although the models retained their individual structural features. The cost-effectiveness of starting treatment at CD4 T-cell counts of 500 cells/mm$^3$ compared with CD4 T-cell counts of 350 cells/mm$^3$ ranged from US$237 to US$1,691 per DALY averted in South Africa and from being cost saving to US$749 per DALY averted in Zambia over a 20-year time horizon (Eaton and others 2014). These estimates were considered likely to be very cost-effective in comparison with international benchmarks for each country and were similar to those in South Africa reported by Alistar, Grant, and Bendavid (2014). However, the threshold for determining whether an intervention is likely to be cost-effective is poorly known in many resource-limited settings. Granich and others (2012) reported lower costs per DALY averted over a shorter time frame of five years and found earlier treatment to be cost-saving over 40 years, using generally more optimistic measures of program quality.

More aggressive public health strategies included treatment at all CD4 T-cell counts and greater expansion of testing and links to the health system. Compared with existing conditions, these scenarios were very cost-effective over 20 years. However, they were less cost-effective over a shorter time horizon, in part because the effect of ART on HIV transmission is initially small but increases (Eaton and others 2014).

Walensky and others (2013) reported on the cost-effectiveness of earlier treatment in a trial-based analysis focused on earlier treatment of serodiscordant couples in South Africa. Treatment of all discordant couples was cost saving in South Africa over five years; over a lifetime, it cost US$590 per life-year saved in South Africa and US$530 per life-year saved in India, both considered very cost-effective. Another trial-based analysis of earlier treatment in serodiscordant couples in Uganda reported a cost per DALY averted of US$1,075 over 10 years (Ying and others 2015).

In the concentrated epidemic setting of India, Maddali and others (2015) and Eaton and others (2014) reported favorable cost-effectiveness ratios for broader strategies of earlier treatment, ranging from US$199 per DALY averted (Eaton and others 2014) to US$512 per quality-adjusted life year gained (Maddali and others 2015). Eaton and others (2014) noted that in the city of Belgaum in southern India, where the epidemic is driven largely by FSWs, the estimated incidence of new infections has fallen substantially since 2003 as a result of programs targeting this special population. They found that the incremental cost-effectiveness ratio for expanding ART to all, regardless of CD4 T-cell counts, was US$131 per DALY averted in the presence of these programs; it was slightly higher (US$241) in the theoretical case in which these programs did not exist. In Vietnam, where the epidemic is driven by FSWs, MSM, and IDUs, the incremental cost per DALY averted was US$289.

To guide resource allocation decisions for different interventions in the presence of budget constraints, Bärnighausen, Bloom, and Humair (2012) reported a favorable incremental cost per infection averted of ART initiated only in persons with CD4 T-cell counts of less than 350 cells/mm$^3$ and VMMC scale-up (US$1,402), compared with early treatment and VMMC scale-up (US$7,325–US$10,083); however, this measure does not value the health- and quality-of-life-related effects of these interventions. This difference was less marked when using the measure of incremental cost-effectiveness of deaths averted (US$7,761–US$10,014 versus US$6,650). In a budgetary analysis, Bärnighausen, Bloom, and Humair (2012) found that VMMC combined with ART at CD4 T-cell counts of less than 350 cells/mm$^3$ was cumulatively less expensive over 12 years and provided a similar incidence reduction as the expansion of treatment eligibility to those with
Table 5.1 Studies Estimating the Cost-Effectiveness of Treatment Direct and Indirect Benefits

<table>
<thead>
<tr>
<th>Study</th>
<th>Study location</th>
<th>Study group</th>
<th>Intervention, comparison</th>
<th>Outcome</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alistar, Grant, and Bendavid 2014</td>
<td>South Africa</td>
<td>HIV-positive adults</td>
<td>ART scale-up at CD4 T-cell level &lt; 350 cells/mm³ vs. test and treat at all CD4 T-cell levels; addition of focused (aimed at highest risk) or general PrEP</td>
<td>At 20 years, cost per DALY gained: US$362–US$370 (all CD4 T-cell levels); US$481–US$486 (&lt; 350 cells/mm³); US$132–US$270 for ART at CD4 T-cell levels &lt; 350 cells/mm³ and focused PrEP; US$1,078–US$1,125 for ART at CD4 T-cell levels &lt; 350 cells/mm³ and general PrEP (2012 US$)</td>
<td>Scale-up of test-and-treat ART strategy is very cost-effective. Focused PrEP strategies, if feasible, would be highly cost-effective combined with ART.</td>
</tr>
<tr>
<td>Bärnighausen, Bloom, and Humair 2012</td>
<td>South Africa</td>
<td>HIV-positive persons; general population</td>
<td>ART at CD4 T-cell level &lt; 350 cells/mm³ vs. at all CD4 T-cell levels; plus VMMC; varying levels of coverage of each intervention</td>
<td>At 12 years, US$1,402 per infection averted for ART (50% coverage) and VMMC (80% coverage); US$7,325–US$10,083 per infection averted for ART at all CD4 T-cell levels (20–80% coverage) and 80% VMMC; US$6,650 per death averted (50% ART at CD4 T-cell levels &lt; 350 cells/mm³); US$7,761–US$10,014 per death averted (70–80% ART at CD4 T-cell levels &lt; 350 cells/mm³; 20–80% ART at all CD4 levels; 80% VMMC) (2012 US$)</td>
<td>Using cost per infection averted as a measure, ART at CD4 T-cell levels &lt; 350 cells/mm³ with VMMC scale-up has lowest cost-effectiveness ratios, whereas ART at all levels has higher ratios.a</td>
</tr>
<tr>
<td>Eaton and others 2014</td>
<td>South Africa; Zambia; Bangalore; Manipur; and Belgaum, India; Vietnam</td>
<td>HIV-positive adults, including key populations</td>
<td>ART at CD4 T-cell levels &lt; 350 cells/mm³ vs. &lt; 500 cells/mm³ and at all CD4 T-cell levels</td>
<td>At 20 years, costs per DALY averted at CD4 T-cell levels &lt; 500 cells/mm³ vs. &lt; 350 cells/mm³: South Africa: US$237–US$1,691; Zambia: dominating to US$749; Vietnam: US$290; at all levels vs. CD4 T-cell levels &lt; 350 cells/mm³: South Africa: US$438–US$3,790; Zambia: dominating; US$790; India: US$131 (all) and US$199 (&lt; 500 cells/mm³); Vietnam: US$289 (all) (2012 US$)</td>
<td>ART at CD4 T-cell levels &lt; 500 cells/mm³ and ART at all levels is very cost-effective in South Africa, in Zambia, and for concentrated epidemic settings in India and Vietnam.</td>
</tr>
<tr>
<td>Granich and others 2012</td>
<td>South Africa</td>
<td>HIV-positive adults</td>
<td>Best-case testing and ART: CD4 T-cell levels &lt; 200 cells/mm³, &lt; 350 cells/mm³, and &lt; 500 cells/mm³ vs. expanded testing and ART at all levels</td>
<td>At five years, ART at CD4 T-cell levels &lt; 500 cells/mm³ vs. 350 cells/mm³: US$221 per DALY averted; ART at all CD4 T-cell levels: US$1,728 per DALY averted; with enhanced prevention (40% reduction in HIV incidence): US$233 for CD4 T-cell levels &lt; 500 cells/mm³, and US$1,767 for all CD4 T-cell levels (2012 US$)</td>
<td>Early ART is very cost-effective when considering a short time frame of five years, with projected cost savings over a 40-year time horizon.</td>
</tr>
<tr>
<td>Hontelez and others 2011</td>
<td>Hlabisa, South Africa</td>
<td>HIV-positive adults</td>
<td>ART at CD4 T-cell levels ≤ 350 cells/mm³ or ≤ 200 cells/mm³</td>
<td>Costs of treating patients at CD4 T-cell levels ≤ 350 cells/mm³ or ≤ 200 cells/mm³ by 2017: breakeven of cumulative net costs in 2026 (2010 US$)</td>
<td>Front-loaded costs of treating at CD4 T-cell levels ≤ 350 cells/mm³ or ≤ 200 cells/mm³ may be offset by model-projected savings from health and prevention gains after 2026.</td>
</tr>
</tbody>
</table>
Table 5.1 Studies Estimating the Cost-Effectiveness of Treatment Direct and Indirect Benefits (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study location</th>
<th>Study group</th>
<th>Intervention, comparison</th>
<th>Outcome</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hontelez and others 2016</td>
<td>10 countries in Sub-Saharan Africa&lt;sup&gt;a&lt;/sup&gt;</td>
<td>HIV-positive adults</td>
<td>ART at all CD4 T-cell levels with continued scale-up vs. no further scale-up</td>
<td>Over 35 years, US$269 per life-year saved</td>
<td>Treatment scale-up at all CD4 T-cell levels was cost-effective.</td>
</tr>
<tr>
<td>Long, Brandeau, and Owens 2010</td>
<td>United States</td>
<td>HIV-positive adults; general population</td>
<td>Expanded ART (75% coverage) vs. status quo, testing and expanded ART</td>
<td>At 20 years, cost per QALY gained for expanded ART: US$21,647; testing (low risk once, high risk annually) and expanded ART: US$22,055; testing (low risk every three years, high risk annually) and expanded ART: US$31,274 (2012 US$)</td>
<td>Expansion of ART coverage in the United States is very cost-effective, and expanded testing increases QALYs gained. Expansion of ART alone decreases new infections by 10.3%; addition of testing (low risk once and high risk annually) decreases new infections by 17.3% over 20 years.</td>
</tr>
<tr>
<td>Maddali and others 2015</td>
<td>India</td>
<td>HIV-positive adults</td>
<td>Early (CD4 T-cell levels ≥ 350 cells/mm³) vs. delayed (CD4 T-cell levels &lt; 350 cells/mm³) initiation of ART</td>
<td>At 20 years, cost per QALY gained: US$442 (ideal program conditions)—US$530 (realistic parameters for program performance) for ART at CD4 T-cell levels ≥ 350 cells/mm³ (2014 US$)</td>
<td>Early treatment is very cost-effective in India.</td>
</tr>
<tr>
<td>Nichols and others 2014</td>
<td>Macha, Zambia</td>
<td>HIV-positive adults; sexually active adults; general population</td>
<td>Early (CD4 T-cell levels &lt; 500 cells/mm³) vs. delayed (CD4 T-cell levels &lt; 350 cells/mm³) ART, with and without PrEP for most sexually active or general population</td>
<td>Cost per QALY gained: US$62 for ART at CD4 T-cell levels &lt; 500 cells/mm³; ART at CD4 T-cell levels &lt; 500 cells/mm³ and general population PrEP: $6,861 (2012 US$)</td>
<td>Early ART is very cost-effective in this rural setting. Adding PrEP was not cost-effective.</td>
</tr>
<tr>
<td>Walensky and others 2013</td>
<td>South Africa; India</td>
<td>HIV-positive partners in serodiscordant couples</td>
<td>Early (CD4 T-cell levels &lt; 550 cells/mm³) vs. delayed (&lt; 350 cells/mm³) initiation of ART</td>
<td>Five-year cost per life-year saved: South Africa: cost saving; India: US$1,800; lifetime horizon, incremental cost per life-year saved: South Africa: US$590; India: US$320 (2012 US$)</td>
<td>Early initiation was cost saving in South Africa over a five-year interval, cost-effective in India at five years, and very cost-effective in both countries over a lifetime.</td>
</tr>
<tr>
<td>Ying and others 2015</td>
<td>Kampala, Uganda</td>
<td>Partners in serodiscordant couples</td>
<td>Scale-up ART at CD4 T-cell levels &lt; 500 cells/mm³ vs. current ART uptake; PrEP for serodiscordant couples</td>
<td>At 10 years: cost per DALY averted: US$1,075 for scale-up ART at CD4 T-cell levels &lt; 500 cells/mm³ (2012 US$); cost per DALY averted: US$1,535 (2012 US$) for ART scale-up at CD4 T-cell levels &lt; 500 cells/mm³ and PrEP</td>
<td>ART scale-up at CD4 T-cell levels &lt; 500 cells/mm³ is very cost-effective compared with current ART uptake. The addition of PrEP to ART scale-up averted more DALYs, but was not cost-effective by per capita GDP standards.</td>
</tr>
</tbody>
</table>


Note: ART = antiretroviral therapy; DALY = disability-adjusted life year; HIV = human immunodeficiency virus; PrEP = preexposure prophylaxis; QALY = quality-adjusted life year; VMNC = voluntary medical male circumcision. This table follows the World Health Organization–endorsed convention, which classifies interventions in cost-effectiveness studies as (1) “cost-effective” when they avert a DALY (or gain a QALY) at a cost of three times per capita GDP or (2) “very cost-effective,” when they avert a DALY (or gain a QALY) at a cost of one times per capita GDP. Dominating = one intervention dominates another intervention when it provides a greater health benefit at a lower cost (Eaton and others 2014, 26).

a. Cost per infection averted does not value ART’s long-term health and health-related quality-of-life effects in the denominator of the cost-effectiveness ratio as recommended by consensus guidelines for cost-effectiveness.
CD4 T-cell counts of more than 350 cells/mm$^3$. Ying and others (2015) reported that scaling up ART to CD4 T-cell levels of less than 500 cells/mm$^3$ among serodiscordant couples in Uganda was very cost-effective compared with status quo uptake at CD4 T-cell levels less than or equal to 350 cells/mm$^3$, whereas the addition of PrEP raised the cost per DALY averted to more than three times the GDP per capita of Uganda. When using a denominator of infections averted that did not include health-related effects, they reported more favorable estimates of cost-effectiveness of the addition of PrEP to ART scale-up. The emerging issue of ART as a component of combination prevention strategies is further examined in chapter 7 of this volume (Garnett and others 2017).

In general, these myriad cost-effectiveness analyses present persuasive model-based projections of the incremental cost-effectiveness of expanding treatment access across a range of settings and populations in LMICs and high-income countries, especially where the costs of outreach are low and a long-term perspective is taken. The following sections consider further methodological issues in modeling the cost-effectiveness of TasP and the impact of TasP on global health recommendations and policy.

**Key Considerations and Limitations of Models Projecting the Cost-Effectiveness of Treatment as Prevention**

As with the modeling of nearly any intervention, many decisions must be made about model structure, and parameters need to be estimated. In general, model parameters can be easily explored in sensitivity analyses, whereas structural choices may have large impacts and their influence may be more difficult to ascertain. In the case of HIV cost-effectiveness, models have been calibrated, to the extent possible, with increasingly good inputs and outputs from real-life implementation of care and treatment programs. Even so, the HIV response in resource-limited settings has only been active for the past 11–12 years, and most modeling horizons are a lifetime; accordingly, certain parameter choices—especially as they relate to evolving drug costs and future lines of treatment—are likely to entail substantial uncertainty.

The best approach to calibrating effectiveness models to underlying HIV epidemic trends has been debated and managed differently by various groups (Hallett, Eaton, and Menzies 2014; Okano and Blower 2014). Modeling cost-effectiveness of TasP also requires examining initiation of treatment earlier in the course of disease than has been done in most LMICs.

The challenge of introducing earlier initiation, along with the valuation of the prevention effects of treatment, is the scarcity of programmatic data (outside of trials and well-studied cohorts) at this stage of the disease to guide parameter development. Many new parameters need to be considered and potentially included, which introduces greater uncertainty (Bärnighausen, Salomon, and Sangrujee 2012).

Substantial uncertainty exists about whether the unit costs of various elements of HIV care and treatment will remain constant as scale-up is accelerated, extended to persons with higher CD4 T-cell counts, or both. As Meyer-Rath and Over (2012) noted, many models rely either on fixed unit costs for a year of treatment or on cost accounting identities, in which each of the constituent costs of treatment is estimated and multiplied by health care utilization figures. These costs may need to be considered in a flexible cost-function manner, given the likely nonlinearities of inputs around scaling TasP, such as uneven need for new infrastructure (Meyer-Rath and Over 2012). Many models also assume constant antiretroviral prices into the future, but for any given drug, prices tend to decline over time (Holmes and others 2010). Conversely, when newer drugs replace existing drugs in guidelines, abrupt price increases may result (Waning and others 2009).

Most cost-effectiveness models further assume that ART retention and adherence do not vary by stage of the disease at initiation, the previous health experience of individuals, and other determinants of health-care-seeking behavior. However, earlier initiation may lead to reduced—or increased—overall retention and adherence, due to underlying differences in those facts. Finally, 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 (Bärnighausen, Salomon, and Sangrujee 2012; Smit and others 2015).

Risk compensation is another consideration that could have a positive or negative effect on projected benefits (Dukers and others 2001; Stolte and others 2004). Changes in HIV risk taking following ART scale-up have at times been considered risk compensation—that is, HIV-negative persons take more sexual risks as ART coverage lowers the average risk of HIV acquisition and ART availability reduces the potential health losses. However, persons could take fewer sexual risks, including changes in sexual behavior, substance use, and contraceptive use, in response to ART scale-up because of improved survival expectations and increased optimism for the future (Bor and others 2013; Raifman and
others 2014). Future models of combination prevention modalities need to attempt to examine the presence of potential behavioral effects.

Select Policy Questions Addressed by Cost-Effectiveness Models
As with nearly all health policy decisions, there are trade-offs among benefits, costs, and risks as national and global decision makers consider possible strategies for their investments in treatment and prevention. Although most TasP scenarios are considered cost-effective or very cost-effective over reasonable horizons for decision making, they all have higher up-front costs. Therefore, optimal policy choices will vary greatly according to available funding, local HIV response, goals, and other elements of feasibility and local preference, such as the political environment. This section highlights several critical policy issues and illustrates the trade-offs involved in using select modeling and cost-effectiveness analyses.

What Trade-Offs Are Involved in Expanding Testing and Links to Care under Existing Treatment Guidelines, Compared with Expanding Earlier Treatment?
Eaton and others (2014) considered the comparative effects on new HIV infections and cost-effectiveness of a policy decision in South Africa, where estimated ART coverage under existing guidelines (CD4 T-cell counts of less than 350 cells/mm$^3$) was approximately 50 percent at the time of the analysis, and in Zambia, where reported ART coverage (CD4 T-cell counts of less than 350 cells/mm$^3$) was more than 90 percent. First, they found that expanding testing and linking patients to care under existing guidelines in South Africa had a substantial effect on lowering the incidence of new infections (6 percent to 28 percent, depending on the model). The approach averted more infections than changing eligibility to CD4 T-cell counts of less than 500 cells/mm$^3$ without expanding testing and links to care (5 percent to 12 percent of infections averted). In Zambia, where reported coverage was already high, simply expanding ART eligibility averted 21 percent to 40 percent of new infections, an impact greater than expanding both testing and links to care (8 percent to 17 percent). Given the high estimated coverage reported in Zambia, the model assumed that cases were being identified earlier and that expanding testing and links to care would have somewhat less of an impact than simply raising the threshold. Since the time of this analysis, Zambia has raised the treatment threshold to CD4 T-cell counts of less than 500 cells/mm$^3$.

From a cost-effectiveness perspective, Zambia’s decision to expand eligibility was generally supported by assessments of costs per DALY averted. The picture for South Africa is less clear, and the conclusions from the models are conflicting. The up-front costs of expanding testing and links to care are high, and four of seven models favored simply expanding eligibility; three favored expanding testing and links to care at the current threshold of CD4 T-cell counts of less than 350 cells/mm$^3$ (Eaton and others 2014). In 2012–13, South Africa chose to expand coverage of persons with CD4 T-cell counts of less than 350 cells/mm$^3$ and to use less toxic antiretroviral medicines; in 2014–15, the country shifted to a policy of treating persons with CD4 T-cell counts of less than 500 cells/mm$^3$; in mid-2016, the government announced the intention to shift to a “treat all” policy consistent with updated WHO guidelines (WHO 2016).

How Might the Timing of the Costs and Benefits of TasP Affect Policy Makers’ Decisions, and What Further Information Could Be Helpful?
Although numerous strategies are considered cost-effective, treating greater numbers of people brings greater up-front costs regardless of potential downstream (discounted) savings. Expanding access to persons with CD4 T-cell counts of less than 350 cells/mm$^3$ results in additional costs that tend to increase over time after a small initial spike (Eaton and others 2014). This increase occurs because increasing numbers of individuals live longer and incur costs to the health system.

In this analysis, increasing eligibility to persons with CD4 T-cell counts of less than 500 cells/mm$^3$, with or without expanded access, results in a greater initial spike in costs; however, as with the “treat all” strategy, the cost curve generally declines each subsequent year, in part because these strategies are expected to reduce the number of new HIV infections (Eaton and others 2014). Accordingly, later outlays could diminish with larger up-front investments, even when including the effects of discounting. Decision making will hinge on the assessment of potential impacts, relevance of model assumptions to the environment, validity of available model inputs, availability of funding and competing investments, and other local factors.

These types of models generally do not consider the financial costs to patients of starting treatment earlier. These costs are related largely to transport and opportunity costs; depending on the extent of decentralization of HIV services, they could be substantial (Rosen and others 2007). If treatment is more for prevention than for direct health benefits, these higher up-front costs may discourage patients from getting care, although this theoretical risk requires further empirical data. Several investigators have demonstrated greater productivity with HIV treatment in clinic- and population-based cohorts (Bor and others 2012), which
is excluded in most model-based analyses. The extent to which productivity gains could offset transport and other costs among people starting treatment while wealthier is unknown.

It is also difficult for models to reflect that, in the context of constrained budgets, additional spending on HIV treatment will lead to displaced resources for other interventions. If the treatment intervention generates health per cost at a rate greater than a certain threshold, then despite that displacement, there is a net gain in health. However, quantifying that threshold is difficult. International guidance has, until recently, suggested benchmarks related to the GDP per capita of a country; there are indications that a more realistic assessment of the opportunity costs of health expenditure would demand lower cost per health gain for an intervention to be likely to be cost-effective (Woods and others 2015).

Policy makers in many sectors, but particularly health, face a trade-off between higher up-front costs and longer-term gains. Unlike problems associated with non-communicable diseases, TasP could reduce the intensity and spread of a transmissible pandemic. In this time-dependent context, donors and policy makers have often leaned toward up-front investments.

What Approaches Could Improve the Cost-Effectiveness of Treatment as Prevention?

Innovations in care delivery, such as task-shifting elements of the delivery cascade to lower-level staff members, have been widely adopted and have facilitated reductions in the costs of delivering care. Innovators such as Médecins Sans Frontières and national governments have further pioneered methods of care delivery, now known as differentiated care, which target the intensity of care to the needs of patients (Duncombe and others 2015; Holmes and Sanne 2015). For example, the formation of community adherence groups among stable patients in Mozambique allowed for substantially less clinic contact and increased retention in care among those opting into these models (Decroo and others 2011). Greater attention to both the models of care and the costs of care delivery is another tool that can be used by in-country stakeholders to strive for greater efficiency and quality of care delivery. The President’s Emergency Plan for AIDS Relief (PEPFAR) program’s expenditure analysis approach entails the collection of detailed data that allow country-level decision makers to distinguish between low- and high-cost providers of quality care. A report of results from this methodology noted reductions in the heterogeneity of the costs of supporting not only ART, but also HIV testing and other key elements along the TasP cascade (PEPFAR 2012).

NATIONAL AND GLOBAL GUIDANCE, POLICIES, AND RESOURCE ALLOCATION

The release in May 2011 of the HPTN 052 data on the remarkable reduction of sexual transmission of HIV disease among serodiscordant couples sent ripples not only through the scientific community, but also through policy-making circles, including national governments, the WHO, and leading funders of HIV programs in low-resource settings (El Sadr and others 2011). After more than two decades of focusing on the health-related effects of ART, experts began to grapple with its role as a tool for the prevention of sexual transmission of HIV.

Members of the WHO Guidelines Group on Couples HIV Testing and Counseling incorporated the HPTN 052 findings in their review process (WHO 2012a), judging them to be directly and immediately applicable to couples counseling and testing. In addition to other potential interventions, the evidence for treatment in this setting was considered to be substantial. After stakeholder reviews, updated guidelines were released 11 months after the HPTN 052 results (box 5.1).

In April 2012, the WHO incorporated the findings of HPTN 052 and released the programmatic update on Option B+ (WHO 2012b). In addition to programmatic data from Malawi that supported full treatment for pregnant women, the potential effects of treatment on sexual prevention of HIV was considered as follows: “If Option B+ can be supported, funded, scaled up at the primary care level and sustained, it will also likely provide the best protection for the mother’s health, and it offers a promising new approach to preventing sexual transmission and new HIV infections in the general population” (WHO 2012b, 4).

In 2013, the WHO combined all of its ART-related HIV guidance into a single guideline that considered the sexual prevention effects of treatment demonstrated in

Box 5.1

Recommendations from the WHO Couples HIV Testing and Counseling Guidelines

People with HIV whose partners do not have HIV and who are started on ART for their own health should be advised that ART is also recommended to reduce HIV transmission to uninfected partners. This is a strong recommendation based on high-quality evidence.
HPTN 052. The guidelines committee changed eligibility criteria to CD4 T-cell counts of 500 cells/mm³, persuaded by a combination of clinical benefits drawn mainly from LMIC settings (Kitahata and others 2009), along with evidence of reduced sexual transmission and tuberculosis among treated individuals. In addition, review of the evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system led to recommendations for treatment regardless of CD4 T-cell count for pregnant women, serodiscordant couples, persons with tuberculosis, and persons with severe liver disease. The U.S. and European guidelines, which already counseled earlier initiation of treatment, also incorporated the benefits of TasP (DHHS 2014; European AIDS Clinical Society 2014). Updated WHO guidelines, reviewed in the setting of further evidence for the health benefits of earlier ART initiation, recommended offering treatment to all individuals with HIV, regardless of CD4 T-cell count (INSIGHT START Study Group 2015; TEMPRANO ANRS Study Group 2015; WHO 2016).

Since the release of the WHO documents recommending earlier treatment, national governments have weighed guidelines changes. As of 2015, the WHO reported that of the 58 WHO focus countries, 34 (59 percent) have adopted new guidelines for treatment of serodiscordant couples and 37 (64 percent) have raised their thresholds to CD4 T-cell counts of 500 cells/mm³. Rwanda’s national guidelines support a TasP strategy (ART regardless of CD4 T-cell count) and universal testing. Other countries with limited health infrastructure and financial resources are struggling with high unmet needs at lower CD4 T-cell counts and have legitimate concerns about crowding out treatment slots for sicker patients (Linas and others 2006). The prevention effects of treatment have also substantially influenced policy and allocation decisions of major payers and funders of HIV programs in LMICs.

The results of HPTN 052 arrived during the latter stages of a global economic downturn and coincided with a leveling off of the rapid growth of HIV-related overseas development aid (Kaiser Family Foundation 2015). Despite these challenges, global leaders, advocates, and public health officials were energized by the potential impact of the addition of ART to the combination prevention armamentarium.

The PEPFAR program is the largest bilateral program supporting the HIV response in LMICs. Its federally chartered Scientific Advisory Board reviewed the HPTN 052 data, discussed its potential applications, and ultimately recommended that PEPFAR support the use of ART in specific populations with CD4 T-cell counts greater than 350 cells/mm³ to prevent transmission to others (El Sadr and others 2011). The board also recommended that “careful evaluations, including assessment of benefit/risk/impact/feasibility and modeling exercises are urgently needed to identify populations that should be prioritized for this intervention, given local conditions” (El Sadr and others 2011, 19). Following these and other recommendations, the U.S. government strongly endorsed accelerating combination prevention in 2011, including a 50 percent increase in persons on treatment over a two-year period. As overall allocations to HIV reached a plateau, PEPFAR weighted its financial allocations more heavily toward treatment and several other high-impact interventions (Cohen, Holmes, and others 2012; Goosby and others 2012; Holmes and others 2012; PEPFAR 2014). In 2014, PEPFAR took the further step of endorsing the UNAIDS (Joint United Nations Programme on HIV/AIDS) 90-90-90 targets, and in 2015 announced support for the ongoing expansion of treatment in the context of further studies on the effectiveness of earlier treatment on individual health (PEPFAR 2014, 2015).

ONGOING RESEARCH

Even as investigators have advanced the understanding of the preventive role of treatment, critical questions remain. Ongoing studies are evaluating the biological, pharmacologic, clinical, and public health elements of ART as a prevention modality, and priority areas for future research continue to emerge.

Biological and Pharmacological Studies

The quest to understand the biology of transmission revolves around the fitness requirements of the viral pathogen (Carlson and others 2014) and its susceptibility to innate host defenses (Borrow 2011). Successful viral suppression does not prevent intermittent viral shedding in the genital tracts of both men (Kalichman and others 2010) and women (Cu-Uvin and others 2010). However, observational studies suggest that the viruses being shed during viral suppression are likely compromised and not readily transmitted (Cohen and others 2013).

Perhaps the most important TasP consideration lies in the simplification of treatment itself. Successful suppression of viral replication virtually eliminates HIV transmission (Muessig and Cohen 2014). Accordingly, linked transmission events reflect failed treatment or resistance to the antiretroviral regimen being used. Failure to adhere to a treatment regimen is the greatest problem. Newer antiviral agents are very well tolerated but still require daily medication.
Glaxo-Smith-Kline and Janssen are exploring a combination of injectable agents, with one combination being tested in the Phase 2b LATTE (Long-Acting Antiretroviral Treatment Enabling) Trial (Margolis and others 2014). The trial includes a run-in of oral agents for safety testing, which has been completed, followed by maintenance injections every month or every two months. Injectable ART may be appropriate for several types of patients, but it is particularly attractive for people in serodiscordant sexual relationships.

**Clinical, Public Health, and Population Effects Studies**

Granich and others (2011) identified more than 50 ongoing studies covering the impact of treatment on prevention among serodiscordant couples and key populations and the secondary benefits of treatment for individuals infected with both tuberculosis and HIV. These studies examine HIV incidence and mortality in the general population and economic outcomes for patients receiving TasP.

Substantial funding has been allocated to research aimed at understanding the potential population-level impact of TasP. Four large studies underway are designed to demonstrate that TasP, as part of a package of prevention interventions, reduces population-level HIV incidence in generalized epidemics (box 5.2).

**CONCLUSIONS**

Treatment has well-known direct effects on the health outcomes of HIV-positive individuals; it has also been conclusively shown in observational studies and randomized clinical trials to prevent sexual transmission of HIV. These prevention effects are backed by years of basic science and clinical work that have established the effectiveness of ART in reducing HIV replication in blood and genital tissues and secretions. The projected effectiveness of early treatment on reductions in sexual transmission of HIV is dependent on the performance of the treatment cascade, including HIV testing and links to care, retention in care, and virological suppression. When the prevention effects of ART are included in analyses of the cost-effectiveness of earlier treatment, ART is generally found to be a highly cost-effective intervention in diverse settings of varied income levels, HIV

---

**Box 5.2**

**Population-Level Effects of Treatment as Prevention: Select Studies**

Population Effects of ART to Reduce HIV Transmission (PopART), also known as the HIV Prevention Trials Network 071, examines the effect of universal testing and treatment compared with treatment at CD4 T-cell counts of less than 350 cells/mm³ or the standard of care. PopART includes 21 communities in two countries—South Africa and Zambia—with a total population of 1.2 million; results are expected in 2017–18 (Hayes and others 2014).

The Botswana Combination Prevention Project assesses the provision of treatment to all individuals with CD4 T-cell counts greater than 350 cells/mm³ or with a viral load greater than or equal to 10,000 copies per milliliter, compared with individuals receiving the standard of care in the setting of scaled-up combination prevention (CDC 2013).

The Africa Centre and the Agence Nationale de Recherche sur le Sida 12249 TasP Trial was developed to establish the causal impact of TasP (treatment as prevention) on population-level HIV incidence and other health, economic, and social outcomes. The trial randomized 34 communities with a total adult population of 34,000 to receive home-based HIV testing and ART referral under either a TasP strategy (intervention) or the South African standard of care with home-based HIV testing (control) (Iwuji and others 2013). The only difference between the intervention and the control arm was whether HIV-positive people were offered immediate ART in early stages of the disease (intervention) vs. only in later disease stages (control).

The SEARCH (Sustainable East Africa Research on Community Health) Study, based in Kenya and Uganda, includes 32 communities of approximately 10,000 individuals each. The study compares early treatment to standard of care, with multiple health outcomes and a prevention outcome of community viral load (Chamie and others 2012; Jain and others 2013).
epidemic types and risk populations. It is projected to be most effective and cost-effective when paired with efforts to identify infected individuals through expanded testing and links to care.

Ongoing population-based studies will provide further information on the wider prevention impact of earlier treatment as part of a package of combination prevention modalities.

ANNEXES

The annexes to this chapter are as follows. They are available at http://www.dcp-3.org/infectiousdiseases.

- Annex 5A. Studies of TasP in Serodiscordant Couples
- Annex 5B. Ecological Studies Examining the Effectiveness of ART on HIV Incidence

NOTES

The authors are indebted to Musonda Namuyemba and Megan Wolf, MPH, for their editorial assistance.

This chapter is linked closely with chapters on the HIV care continuum in adults and children (chapter 4, Harrispersaud and others 2017), prevention of mother-to-child transmission (chapter 6, John-Stewart and others 2017), and cost-effectiveness of interventions to prevent HIV acquisition (chapter 7, Garnett and others 2017).

World Bank Income Classifications as of July 2014 are as follows, based on estimates of gross national income (GNI) per capita for 2013:

- Low-income countries (LICs) = US$1,045 or less
- Middle-income countries (MICs) are subdivided:
  (a) lower-middle-income = US$1,046 to US$4,125
  (b) upper-middle-income (UMICs) = US$4,126 to US$12,745
- High-income countries (HICs) = US$12,746 or more.

REFERENCES


Lu, W., G. Zeng, J. Luo, S. Duo, G. Xing, and others. 2010. “HIV Transmission Risk among Serodiscordant Couples: A Retrospective Study of Former Plasma Donors in...


