The spread, treatment, and prevention of HIV-1: evolution of a global pandemic

Myron S. Cohen, …, Kevin DeCock, Joep Lange


The most up-to-date estimates demonstrate very heterogeneous spread of HIV-1, and more than 30 million people are now living with HIV-1 infection, most of them in sub-Saharan Africa. The efficiency of transmission of HIV-1 depends primarily on the concentration of the virus in the infectious host. Although treatment with antiviral agents has proven a very effective way to improve the health and survival of infected individuals, as we discuss here, the epidemic will continue to grow unless greatly improved prevention strategies can be developed and implemented. No prophylactic vaccine is on the horizon. However, several behavioral and structural strategies have made a difference — male circumcision provides substantial protection from sexually transmitted diseases, including HIV-1, and the application of antiretroviral agents for prevention holds great promise.
The spread, treatment, and prevention of HIV-1: evolution of a global pandemic

Myron S. Cohen,1 Nick Hellmann,2 Jay A. Levy,3 Kevin DeCock,4 and Joep Lange5


The most up-to-date estimates demonstrate very heterogeneous spread of HIV-1, and more than 30 million people are now living with HIV-1 infection, most of them in sub-Saharan Africa. The efficiency of transmission of HIV-1 depends primarily on the concentration of the virus in the infectious host. Although treatment with antiviral agents has proven a very effective way to improve the health and survival of infected individuals, as we discuss here, the epidemic will continue to grow unless greatly improved prevention strategies can be developed and implemented. No prophylactic vaccine is on the horizon. However, several behavioral and structural strategies have made a difference — male circumcision provides substantial protection from sexually transmitted diseases, including HIV-1, and the application of antiretroviral agents for prevention holds great promise.

Introduction

Now, 28 years after AIDS was first recognized (1), HIV-1 requires continued global focus and investment. The cumulative total of individuals infected with HIV-1 and deaths due to AIDS since the pandemic began exceeds 60 million and 25 million people, respectively (2). At the end of 2007, the Joint United Nations Programme on HIV/AIDS (UNAIDS) and the WHO estimated that there were 33.2 million people living with HIV-1 (see Global estimates of the number of individuals affected by the HIV/AIDS pandemic in 2007), that 2.5 million individuals became newly infected with HIV-1 in 2007, and that 2.1 million people died of AIDS in that year (3). Analysis of the most recent global data confirms the disproportionate impact of HIV/AIDS on sub-Saharan Africa, however, HIV-1 epidemics are concentrated in the Americas, Europe, and Australia is clade B, whereas clade C predominates in the most heavily affected part of the world, southern Africa. Increasingly, recombinant HIV-1 (recombining different clades) are detected: A/G recombinant strains are prevalent in West Africa, and B/C recombinant strains are prevalent in China (8). As will be discussed below, different clades might be transmitted with different levels of efficiency and might differ in their pathogenic potential.

In comparison with HIV-1, HIV-2 is much less prevalent, and individuals infected with HIV-2 are primarily found in West Africa and India and, to a more limited extent, in Portugal and former Portuguese colonies. In addition, infection with HIV-2 is associated with a slower progression to immune deficiency, and the virus seems to be less efficiently transmitted, even from infected women to their offspring (9). Although HIV-2 has no clades, there are 5 groups, of which A and B are the most prominent (9, 10). To date, differences between HIV-1 and HIV-2 have not led to major insights that have improved treatment or prevention of infection with HIV-1. Due to space constraints, we restrict our discussion to HIV-1 unless otherwise noted.

In summary, over the past 28 years HIV-1 has moved from a single report of an infection cluster (1) to a worldwide pandemic (3). In this Review, we outline the unequal distribution of the disease across the planet and the biological and behavioral factors that have led to the current situation. We also discuss the molecular pathogenesis of infection, including the unique interaction between the virus and host cells that leads to immunodeficiency and death. Finally, we describe the evolution of HIV-1 treatment

Nonstandard abbreviations used: HAART, highly active antiretroviral therapy; HIV-2, herpes simplex virus-2; IRIS, immune reconstitution inflammatory syndrome; LTNP, long-term nonprogressor; MSM, men who have sex with men; PRED, plasmacytoid DC; STD, sexually transmitted disease; TB, tuberculosis.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: J Clin Invest. 118:1244–1254 (2008). doi:10.1172/JCI34706.
The epidemiology of HIV/AIDS

Perhaps one of the most surprising aspects of the HIV/AIDS pandemic is the unequal spread of HIV-1. Although no population has been spared (HIV-1 does not respect social status or borders), some regions and populations have been affected far more than others. Unequal spread of HIV-1 reflects a broad combination of biological and social factors.

HIV-1 can be transmitted by contaminated blood and blood products (most importantly by injection drug usage), from an infected mother to her baby (before, during, or after birth and through breast milk), and, most frequently, through either vaginal or anal intercourse. In this section, we depict the geography of HIV-1 by region.

The industrialized world. AIDS was first recognized in the US (1). The US remains the country most heavily affected by the HIV/AIDS pandemic in the industrialized world, and the US epidemic remains a paradigm of HIV/AIDS in the developed world. Currently about three-quarters of newly reported cases of infection with HIV-1 in the US are in men, most of them MSM and especially African American MSM (4). However, there has also been a steady and disturbing increase in the number of women infected with HIV-1 through heterosexual contact (4). Mother-to-child transmission of HIV-1 has been nearly eliminated in the US through routine prenatal screening linked to provision of antiretroviral therapy to those pregnant women who are infected with HIV-1 (11).

In the US, racial and ethnic minorities, especially African Americans and Hispanics, are disproportionately affected by HIV/AIDS, and there is extreme geographic heterogeneity across the country (4). Although the absolute number of cases of infection with HIV-1 are greatest in urban centers, the rural southeastern US is an area of relatively high HIV-1 incidence and prevalence, most probably reflecting a variety of adverse social factors in this region (4).

The HIV-1 epidemic in Western Europe is diverse but was initially fueled by infections among MSM and injecting drug users, the latter especially in the southern part of the continent (3). Italy, Spain, Portugal, France, and the United Kingdom have been most heavily affected (3). Heterosexual transmission of HIV-1 in Europe has slowly increased, and many infections today are found among immigrants from sub-Saharan Africa (3). In Eastern Europe, where brisk and severe epidemics emerged among injecting drug users in the late 1990s, the most affected countries are the Russian Federation and Ukraine (3).

Low- and middle-income countries. Sub-Saharan Africa now accounts for 68% of persons living with HIV/AIDS worldwide and for 76% of all deaths due to AIDS (2, 3). In 8 countries in southern Africa, the prevalence of HIV-1 infection in the general population exceeds 15% (3), and these countries account for almost one-third of the new HIV-1 infections and AIDS deaths in the world. South Africa alone has over 5 million persons infected with HIV-1 (i.e., more than 10% of the population), more than any other country in the world (3, 12).

As of 2007, approximately 61% of individuals infected with HIV-1 in Africa were women (3). Almost 90% of children who are infected with HIV-1 live in Africa (3). Although declines in HIV-1 prevalence have occurred in a number of countries in recent times (most notably in Kenya and Zimbabwe), Africa has witnessed the full devastation of the HIV/AIDS pandemic in a way not seen elsewhere. Important consequences include family disruption, socioeconomic damage, a decrease in life expectancy, a secondary epidemic of tuberculosis (TB), and a dramatic increase in the number of orphans.

The most affected region of the world outside Africa is the Caribbean, where the overall population prevalence of infection with HIV-1 is about 1%, primarily resulting from heterosexual transmission (3). Haiti has been the most heavily affected country, but its HIV-1 prevalence has declined in recent time. HIV/AIDS in Latin America is diverse in distribution and epidemiology, and approximately one-third of persons in Latin America who are infected with HIV-1 live in Brazil. MSM are the group most affected throughout this region.

In South and Southeast Asia, HIV/AIDS has been concentrated in specific groups at risk, including injecting drug users, female commercial sex workers and their clients, and MSM. In countries with a large population, such as India and the People’s Republic of China, concentrated epidemics in specific groups, and low overall prevalence, still convert into large numbers of people infected with HIV-1 (3).

HIV-1 transmission

Transmission of HIV-1 by whatever route depends on the infectiousness of the “index case” (i.e., the person who transmits the HIV-1 virus) and the susceptibility of the naive host (13). Infectiousness depends on the concentration of HIV-1 and HIV-1–infected cells in relevant body fluid (blood or genital tract secretions) (14) as well as virus-specific determinants required for HIV-1 transmission. Biologic strategies for preventing the transmission of HIV-1 to a naive host depend on complete dissection of these rapid and complex biological events.

The concentration of HIV-1 and HIV-1–infected cells in the body. The concentration of HIV-1 in blood and genital secretions varies dramatically depending on the stage of the disease (15). The highest viral loads are detected in the first weeks after infection and in people with advanced disease (Figure 2). The period of acute HIV-1 infection, the first days after infection, can be operationally defined as the window of time following exposure to HIV-1 and the inception of infection and viremia but before HIV-1–specific antibody-mediated immune responses can be reliably detected.
The intense viral replication observed during acute infection (often called ramp-up viremia) is associated with and may be in part caused by a host "cytokine storm" (17). A substantial amount of HIV-1 transmission seems to result from sexual exposure to subjects with acute infection (18). Using empirical data from a study in Uganda (18), Wawer and colleagues reported that HIV-1 transmission during the first 5 months after exposure (which includes acute and early HIV-1 infection periods) was 0.0081 per coital act, compared to 0.0010 and 0.0043 per coital act in people with established or late HIV-1 infection, respectively. In addition, the magnitude of peak viremia in subjects in sub-Saharan Africa infected with HIV-1 clade C (15) generally exceeds that observed in subjects in the US infected with HIV-1 clade B (19), although the basis for these differences (viral or host factors) is not known. Nevertheless, the high viral loads observed at peak viremia in individuals in sub-Saharan Africa infected with HIV-1 clade C can only serve to further the transmission of the virus.

As acute HIV-1 infection resolves, the viral load reaches a steady-state "set point" that must reflect some combination of virus-specific properties, host genetics, and host defenses (20). Additionally, coinfections with other microbial pathogens seem to affect blood and genital tract viral loads (13). For example, modest blood viral load increases have been observed in individuals who are coinfected with the malaria-causing pathogen, which may in turn affect HIV-1 transmission rates and the magnitude of the HIV-1 epidemic in regions where malaria is prevalent (21). Likewise, STDs (both genital ulcers and urethral inflammation) cause reversible increases in cutaneous and genital tract viral concentrations and have consistently been associated with higher prevalence of HIV-1 infection in all geographic regions (13).

Several empirical studies have linked the magnitude of the blood viral burden with HIV-1 transmission (22, 23). Modeling exercises suggest that anything that increases viral burden in the genital tract will increase the efficiency of HIV-1 transmission (13, 24). Not surprisingly, overwhelming numbers of epidemiologic studies have linked HIV-1 and STDs (13, 25), since STDs generally increase the concentration of HIV-1 in the genital tract. In one intervention trial, the treatment of STDs reduced the incidence of HIV-1 (26). Among the STD pathogens, herpes simplex virus–2 (HSV-2) seems to play the most important role because of high worldwide prevalence of infection with the pathogen, recurrent clinical and subclinical HSV-2 reactivation, and concomitant genital ulcers and inflammation (27).

**Virus-specific determinants**. The viral and cellular requirements for HIV-1 transmission remain incompletely understood. Most recently, Brass et al. (28) applied a unique approach, using siRNA to identify "HIV dependency factors." In this work, the authors confirmed 36 well-recognized factors such as CD4 and CXCR4.
However, they also identified more than 200 novel factors that provide unique insight into the intracellular biology of HIV-1.

More than 80% of the HIV-1 variants detected in an individual during acute infection come from a single transmitted isolate (M.S. Cohen, unpublished observations). However, some investigators have reported that women show greater viral diversity than men in the early period after transmission (29), perhaps reflecting the large transmission surface area of the cervix and vagina, the complex biological environment of the vagina, and/or the impact of hormonal contraception in some women. HIV-1 isolates can be distinguished by the coreceptor used to infect cells (30). Virions utilizing CCR5 as a coreceptor (R5 viruses) are macrophage-tropic (30, 31). Those using CXCR4 (X4 viruses) are T cell–tropic and perhaps more cytopathic. Most of the viruses observed soon after transmission (97%) utilize CCR5 coreceptors (ref. 31 and M.S. Cohen, unpublished observations), but the reason for this dominance remains unclear. It is possible that one or more characteristic of R5 viruses are required to infect a mucosal cell or for initial replication, although, it is possible that X4 viruses initiate infection and are then eliminated (32).

Human susceptibility. Human susceptibility to infection with HIV-1 is not uniform. Resistance to infection reflects some combination of genetic factors, innate resistance, and (perhaps very rarely) acquired resistance. Genetic studies have demonstrated that a deletion mutation in the gene encoding CCR5 offers substantial protection from HIV-1 infection (33), and other genetic factors with less powerful impact on acquisition of HIV-1 infection have been reported (34). In the latter case, the gene products that confer resistance to an individual and the way in which this is accomplished are not understood.

Innate resistance might reflect the effects of autochthonous microbial flora and “at-the-ready” immunologic mucosal defenses. For example, women with the change in vaginal flora described to infection (39). However, HIV-1 superinfection seems to occur in other population-based studies have linked either of these defense mechanisms to protection from infection in humans. In some high-risk uninfected subjects, an innate nontyctotoxic HIV-1–specific CD8+ T cell immune response has been correlated with resistance to infection (39), however, HIV-1 superinfection occurs to occur fairly frequently (40), suggesting that infection does not induce protection. Finally, although vaccines that evoke a cell-mediated immune response reduce both peak and early viral load in the rhesus macaque model of infection with HIV-1 (41), to date they have failed to prevent infection in animals or humans (see below).

HIV-1 pathogenesis

After the HIV-1 acquisition event, the interaction between the virus and host determines the natural history of the disease (this varies considerably from person to person) and the probability of a secondary acquisition event(s), which is greatly dependent on the viral load. To date, the destructive properties of HIV-1 have not been completely unraveled. Yet such understanding is required to prolong host survival and to move toward the ultimate goal of curing this infection.

First, it is useful to compare HIV-1 and HIV-2 (42). Although they are both retroviruses, the pathogenesis of the disease that they cause differs dramatically (9). HIV-1 and HIV-2 have similar genetic structures, but they can differ by up to 40% at the DNA sequence level (Figure 3) (42). Both viruses have structural and accessory genes that influence replication and pathogenesis (Table 1). Yet infection with HIV-2 is far less pathogenic than is infection with HIV-1 (9).

Differences in biologic properties in the various group M HIV-1 clades have also been observed. Clade D HIV-1 leads to faster disease progression than the other clades, especially clade A (43). Clade C HIV-1 seems to be spreading more rapidly worldwide than other clades (8). Clade C HIV-1 retains the CCR5-tropic phenotype even during disease progression (44), and this characteristic might provide a transmission advantage if the R5 phenotype is favored (see above).

Attachment of HIV-1 to cells expressing CD4 is followed by binding to viral coreceptors and fusion of the cell and the virion, which is mediated through the viral glycoprotein gp41 (30, 31). After the interaction at the cell surface, the HIV-1 capsid enters the cell, and this is followed by reverse transcription of viral RNA into double-stranded DNA and then integration of viral DNA into the host genome. The integrated virus can produce virions or remain in a latent (silent) state (45). Latent infection might reflect methylation of the viral genome, reduced Tat or Rev expression, or the activity of cellular proteins such as histone deacetylases (HDACs), PTEFb, and/or a CD8+ T cell antiviral factor(s) (known as CAF) (45, 46). Although transmission of HIV-1 generally results from R5 viruses, even when the host is exposed to X4 and R5 variants, after several years, X4 viruses can be detected in at least 50% of individuals infected with HIV-1, and this shift is strongly associated with more rapid CD4+ T cell decline and enhanced progression to AIDS (47, 48). The mechanisms by which X4 viruses arise is not known, nor is the precise relationship of this viral phenotype to falling CD4+ T cell numbers.

CD4+ T cell loss in individuals infected with HIV-1 occurs through direct infection of the CD4+ T cells and through indirect effects resulting from immune activation and apoptosis (49). Loss of gut CD4+ T cells is particularly striking during acute and early HIV-1 infection (50, 51). Particular viral proteins and cytokines are toxic to CD4+ T cells, and some CD8+ T cells kill bystander uninfected CD4+ T cells (49). In addition, the virus might inhibit thymic generation of CD4+ T cell numbers, thereby limiting recovery of CD4+ T cells in the infected host (49).

Differences in virus replication within various cell types can be influenced by intracellular factors that affect the replicative cycle. Among these is the anti–HIV-1 effect of apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G), a cytidine deaminase that produces an inactive proviral DNA product during reverse transcription (52). The action of this enzyme can be blocked by the viral protein Vif, which targets APOBEC3G for degradation (52). Another identified intracellular antiviral factor, tripartite motif–containing 5-α (TRIM5-α), regulates the ability of different viruses to infect human and monkey cells (53). This protein seems to interact with the viral capsid and block uncoating, which is required during an early replicative step (53).

The clinical manifestation of infection with HIV-1 results from the destruction of CD4+ T cells to the point where unusual infectious agents are able to take hold, leading to the “opportunistically” infections that characterize AIDS. However, progression from
Infection with HIV-1 to AIDS is not inevitable, and some people tolerate infection with the virus for long periods of time without health consequences (54, 55). Long-term survivors, also known as long-term nonprogressors (LTNPs), maintain normal CD4+ T cell counts and low viral load in the absence of antiretroviral therapy (54, 55). Such a phenomenon must reflect differences in the HIV-1 variants (e.g., absence of the nef protein; ref. 56), unique host genetic factors (e.g., “protective” HLA polymorphisms; ref. 57), unique host defenses, or all three. Accordingly, it is widely believed that LTNPs can provide clues to our understanding of pathogenesis. For example, several genetic factors seem to regulate viral set point (M.S. Cohen, unpublished observations).

Innate and adaptive immunity have central importance in virus control and have both humoral and cellular immune activities. The primary difference between the innate and adaptive immune system is the quick response of innate immunity to an incoming pathogen. The innate immune system recognizes conformational patterns of pathogens rather than specific epitopes (58). Soluble factors involved in innate immune responses include mannose-binding lectins and complement, both of which can inactivate HIV-1 directly (59).

Among the cellular immune participants are DCs, NK cells, and a special subset of CD8+ T cells. DCs play a role in both innate and adaptive immunity. Various DC types found throughout the body have the major function of presenting antigen to T cells and B cells, resulting in efficient adaptive immune responses (60). The plasmacytoid DCs (pDCs), the dominant producers of type 1 IFNs (60–62), are found in CD4-rich regions of lymphoid tissues and in low numbers in the blood. PDC levels are reduced during infection with HIV-1, particularly in individuals who have progressed to AIDS, but are higher in LTNPs than in healthy controls (63), suggesting that PDCs might play a role in restricting disease progression. PDCs release IFN-α in response to HIV-1 and HIV-1–infected cells (64).

NK cells have an important role in destroying HIV-1–infected cells, and their function depends on efficient production of cytokines such as IL-12 and type 1 IFNs (65). CD8+ T cells have been identified that also participate in innate responses. They show anti–HIV-1 activity without killing the infected cell. This CD8+ T cell noncytotoxic antiviral response is mediated by a novel, but undefined, CD8+ T cell antiviral factor known as CAF that blocks viral transcription (46). Such responses are highest in LTNPs, and a decrease in these responses leads to HIV-1 replication and disease progression (54, 55).

Antibody production is the primary humoral antimicrobial immune response of adaptive immunity and seems to be most important during the early stages of infection with most microbes. Antibodies that can neutralize HIV-1 virions are directed at various regions on the HIV envelope glycoproteins gp120 and gp41, and differences in neutralization sensitivity have been demonstrated among HIV-1 isolates (66). Broadly neutralizing antibodies (i.e., antibodies able to inactivate viruses from various clades and groups) can be identified in the serum of a small number of individuals who are infected with HIV-1 (66), but the relevance of these antibodies is not understood. Genetic and nongenetic factors influence viral sensitivity to neutralization; these include the expression of linear and conformational epitopes, the degree of envelope glycosylation, envelope stability, and whether cellular proteins become attached to the virion surface at the time of budding. Other antibodies that can have a beneficial effect are those that attach to virus-infected cells and induce direct cytotoxic responses of neutrophils and NK cells. Antibody-directed cellular cytotoxicity (ADCC) has been identified in some LTNPs and healthy infected individuals (67). However, some antibodies can bring HIV-1 into T cells and macrophages via the Fc or complement receptor (68) and thereby actually enhance infection; antibodies with these properties have been found in individuals who advance to AIDS (68).

Very early following infection with HIV-1, presentation of antigens by DCs leads to the development of HIV-1–specific CD4+ T cells. These cells can sometimes have cytotoxic activity but primar-
ily help in the development of HIV-1–specific CD8+ T cells and B cells that produce HIV-1–specific antibodies. Activated CD4+ T cells are more susceptible to infection with HIV-1 than other CD4+ T cells in the infected host (69), so destruction of these CD4+ T cells early in infection might cause a lack of immune control of the HIV-1 infection (69).

CD8+ T cell cytotoxic activities are associated with the control of many viral infections (70). This places an emphasis on HIV-1–specific CTL responses in preventing infection and advancement to disease (71). Nevertheless, AIDS patients have HIV-1–specific CTLs (72), and some of these CTLs can be detrimental by lysing uninfected “bystander” CD4+ T cells and DCs (73). Importantly, HIV-1–specific CTLs might not be able to kill virus-infected cells because they lack perforin or other cytotoxic proteins (74). Moreover, HIV-1 has several mechanisms to resist HIV-1–specific CTLs; for example, mutations arise in viral antigenic peptides; DCs express nonimmunogenic viral peptides; and HIV-1–infected APCs exhibit defects in antigen presentation.

**Treatment for individuals infected with HIV-1**

The early years of the HIV/AIDS pandemic were characterized by a steep learning curve that connected CD4+ T cell numbers (at that time an entirely novel host defense mechanism), viral burden (originally characterized by culture), and characteristic opportunistic infections, reflecting the central importance of CD4+ T cells in the immune response. Improvements in managing opportunistic infections notwithstanding, infection with HIV-1 was usually fatal months or years after the appearance of AIDS-defining symptoms.

The recognition that several different classes of agents working at different places in the viral life cycle could inhibit HIV-1 replication remains the most important advance in this field, for both treatment and prevention. The challenge in global health has been in getting the right antiretroviral drugs to the right people at the right time.

**Antiretroviral treatment.** Antiretroviral drugs are suppressive; they cannot cure infection with HIV-1. However, antiretroviral therapy has changed infection with HIV-1 from an almost uniformly fatal infection into a chronic disease. Indeed, the widespread use of antiretroviral therapy in the US led to rapid and remarkable declines in morbidity and mortality (75). However, shortly after the success of the first agent, azidothymidine (AZT), it became clear that the ability of HIV-1 to generate drug-resistant mutants meant that therapy would require a combination of agents, ultimately referred to as HAART (highly active antiretroviral therapy; ref. 76). HAART combinations are designed to slow the emergence of drug-resistant mutants and permit more durable viral suppression; but they also complicate treatment. In addition, discontinuation of HAART leads to rapid loss of viral suppression (77) as well as other adverse health consequences, including increased risk of opportunistic infections, cardiovascular complications, and death (78).

Despite the success of HAART in the US and Western Europe during the last years of the 20th century, there was initially little push to extend the use of HAART to resource-poor settings where the epidemic was much larger. The 2000 International AIDS Conference in Durban, South Africa, accelerated a breakthrough in worldwide access to HAART. Shortly before the Durban conference, UNAIDS and five major pharmaceutical companies announced an agreement on substantial reductions in the price of the antiretroviral drugs used in HAART combinations for the least-developed nations of the world, especially those in sub-Saharan Africa (the Accelerating Access Initiative). This agreement was followed, in 2001, by a United Nations General Assembly Special Session (UNGASS) on HIV/AIDS that clearly stated that antiretroviral treatment was an essential component of the fight against HIV/AIDS. The WHO subsequently included antiretrovirals on the Essential Medicines List and formulated guidelines for the development of a public health approach to treating individuals in resource-poor settings infected with HIV-1 (79).

More recently, new and substantial funding mechanisms have been established to offer antiretroviral therapy in resource-constrained settings. These include the World Bank’s Multi-Country HIV/AIDS Program (MAP), the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM), and the US President’s Emergency Plan for AIDS Relief (PEPFAR). Through its “3 by 5” initiative, the WHO set the ambitious target of having 3 million people in resource-poor settings on HAART by the end of 2005 (80). Although prices of antiretroviral drugs continued to fall, in part because of generic competition, only 1.3 million people were receiving HAART at the end of 2005 and 2.0 million by the end of 2006 (81). Despite not reaching the target of the 3 by 5 initiative, the scale-up in Africa has been very dramatic, rising from 100,000 people receiving HAART at the end of 2003 to 810,000 at the end of 2005 (81) and 1,340,000 (28% of those in immediate need) at the end of 2006 (81).

Formidable obstacles remain to universal access to antiretroviral therapy (reviewed in ref. 82). In high-prevalence countries, the HIV/AIDS epidemic not only has greatly increased demands on an already understaffed and malfunctioning health sector, but has further compromised the work force because many health care workers are infected with HIV-1. The shortages of doctors and nurses prompt the involvement of less-qualified health care workers and even community members in the delivery of care, so strategies must be clear and simple. Community-based care has been successfully pioneered in diverse, challenging settings, including rural areas and urban slums (83, 84) and deserves to be further explored in other regions.

**Constraints to antiretroviral treatment in resource-poor settings.** The management of HIV/AIDS varies widely by region, depending on financial resources and technology. For example, in high- and middle-income countries, patient information always includes a CD4+ T cell count (reflecting the magnitude of disease progression), the blood plasma viral load (which correlates with the risk and speed of progression), and even viral resistance testing (used to tailor initial therapy and to modify therapy in the event of failure). However, in resource-poor settings, even CD4+ T cell enumeration is not always available, meaning that the decision to initiate antiretroviral therapy is based purely on clinical criteria. Therefore, antiretroviral treatment is often initiated too late in the course of infection to prevent high mortality rates (85). In addition, patients started late on therapy are more likely to suffer an immune reconstitution inflammatory syndrome (IRIS) (86). Although IRIS is incompletely understood, it seems to occur during early restoration of host immune defenses due to HAART in some individuals with chronic opportunistic infections (e.g., tuberculous and non-tuberculous mycobacteria, and fungi) who had been relatively asymptomatic despite low CD4+ T cell numbers and advanced HIV disease.

In resource-constrained countries, the ability to measure plasma HIV-1 RNA load (pVL) to monitor the response of an individual to treatment is even more limited than the ability to enumerate CD4+ T cell numbers. Treatment failure is usually diagnosed late, based on declining numbers of CD4+ T cells and/or clinical disease progression. Since these are usually preceded weeks or months ear-
Overlapping drug toxicities (94) and pharmacological interactions between anti-TB drugs and antiretroviral agents narrow antiretroviral treatment choices in those who need concomitant treatment (95). The question of optimal initiation of anti-TB and antiretroviral therapy in coinfected patients remains a research priority. Last, successful global control of TB very much depends on our ability to prevent and treat HIV-1 infections.

**Constraints to therapy.** The specific antiretroviral therapy choices have been affected by a complex constellation of considerations, and drug rationing has proven necessary, even in the US. Both cost and fear of drug toxicity have limited usage of HAART, but recognition of the advantages of earlier therapy have led to new WHO guidelines suggesting that therapy be initiated when CD4+ T cell numbers fall below 350 cells/mm3 (81). Cost considerations have also led to the widespread use of a generic fixed-dose HAART combination of d4T/3TC/nevirapine (Trionum, Cipla). Although the short-term tolerance of this combination is generally good, after a few years of treatment, many people develop a disfiguring syndrome, which is referred to as lipodystrophy and is primarily ascribed to d4T (96). In addition, HIV-1 resistance to d4T is rapidly emerging in many countries (87).

Prevention of HIV-1 transmission
Approximately six new individuals became infected with HIV-1 for each person started on HAART in 2006 (3, 100). In the absence of curative therapy, control of the HIV/AIDS epidemic requires broad implementation of effective and sustainable prevention measures. Although successful HAART seems to reduce transmission of HIV-1 (101), other methods to prevent transmission are required, because most individuals infected with HIV-1 are unaware of their status, and few actually receive HAART. Furthermore, since no single intervention has been found to be universally effective at preventing the transmission of HIV-1, packages of specific interventions must be designed for geographic settings according to the local transmission patterns and risk behaviors (Figure 1) (102, 103). Currently, however, less than one in five persons at substan-

![Figure 4: Opportunities for preventing infection with HIV-1. Prevention efforts can be divided into four broad categories: those offered to HIV-1-negative subjects, especially those in high-risk groups; those offered to people with a strong likelihood of recent exposure to HIV-1; those offered shortly after exposure; and those offered to people who are already infected.](image)

**Table 2**
Ongoing clinical trials to assess the effectiveness of orally administered antiretroviral drugs in preventing the acquisition of infection with HIV-1

<table>
<thead>
<tr>
<th>Product</th>
<th>Population enrolled in trial</th>
<th>Number of participants</th>
<th>Site of trial</th>
<th>Trial sponsor</th>
<th>Year results expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenofovir</td>
<td>Male and female injecting drug users</td>
<td>2,000</td>
<td>Thailand</td>
<td>CDC</td>
<td>2008</td>
</tr>
<tr>
<td>Tenofovir</td>
<td>MSM</td>
<td>400</td>
<td>US</td>
<td>CDC</td>
<td>2009</td>
</tr>
<tr>
<td>Truvada</td>
<td>Heterosexual men and women</td>
<td>1,200</td>
<td>Botswana</td>
<td>CDC</td>
<td>2010</td>
</tr>
<tr>
<td>Truvada</td>
<td>MSM</td>
<td>3,000</td>
<td>South America, US, South Africa, and Thailand</td>
<td>NIH</td>
<td>2011</td>
</tr>
</tbody>
</table>

CDC, Centers for Disease Control and Prevention.
tial risk of infection with HIV-1 in the developing world has access to basic HIV-1 prevention services (81, 82).

Prevention efforts can be divided into four broad categories: those offered to HIV–1–negative subjects, especially those in high-risk groups; those offered to people with strong likelihood of exposure to HIV-1 but before such exposure; those offered to people shortly after exposure; and those offered to people who are already infected (secondary prevention) (Figure 4). In addition, effective prevention strategies have been developed to reduce mother-to-child transmission of HIV-1 infection (104) and to reduce blood product transmission of HIV-1 through use of sensitive and reliable screening methods (105).

Behavioral prevention methods. Behavioral interventions were the first prevention methods to be broadly implemented. These interventions have focused on sexual abstinence (106), delayed sexual debut, reduced numbers of sexual partners, routine condom use, and reduced needle sharing or clean needle use among injection drug users (107). However, studies of some single behavioral interventions, such as abstinence-only programs in high-income countries, have failed to show substantial benefit on HIV-1 risk behaviors or biological outcomes (108). Reductions in HIV-1 prevalence in many countries, including the US, Australia, Brazil, Thailand, Uganda, Kenya, and Zimbabwe, have been attributed to combinations of behavioral intervention strategies (108).

Male circumcision. Biological interventions for HIV–1–negative subjects hold considerable promise (109). Three recent randomized, controlled clinical trials have demonstrated the efficacy of adult male circumcision in reducing female-to-male transmission of HIV-1 by approximately 50%–60% (110–112). Conversely (and perhaps surprisingly), cervical diaphragms did not offer women significant protection from infection with HIV-1 (113); however, use of diaphragms alone, without concomitant male condom use, was not evaluated. In general, these results offer insight into the acquisition of infection with HIV-1: for men, it seems clear that the foreskin, especially the inner mucosal surface, is a key target for HIV-1 (36); for women, the endocervix is probably not the only site for HIV-1 acquisition (113).

Treating STDs. Although, as discussed above, STDs are well described cofactors in transmission of HIV-1, rendering HIV–1–negative subjects more susceptible to infection (13), clinical studies designed to use treatment for an STD to prevent infection with HIV-1 have provided inconsistent results (25, 114–118). The differences between these trials have been discussed exhaustively (117), and the results emphasize that to achieve prevention of infection with HIV-1 benefit(s) from the treatment of classical ulcerative and mucosal STDs, it is critical to treat the right people and the right STDs at the right point in time. For example, symptomatic and asymptomatic genital HSV-2 infections are believed to greatly contribute to the spread of HIV-1 (13, 27). However, use of HSV-suppressive therapy in HIV–1–negative, HSV-2–positive individuals failed to demonstrate a significant benefit in two large clinical trials (118, 119). Another study nearing completion is designed to determine whether chronic suppressive therapy for HSV-2 ulcers in HIV-1/HSV-2 coinfected individuals reduces transmission of HIV-1 to the HIV–1–negative sexual partner (120).

Vaccines. A vaccine that protects individuals from infection with HIV-1 is required to stop the epidemic, and so an absolute worldwide commitment to this goal has been established (121). However, attempts to develop a protective vaccine have been severely compromised by an incomplete understanding of protective HIV-1 immunity and the inability to induce a sufficiently potent immune response against HIV-1 (122, 123). To date, two vaccine candidates, a gp120 envelope glycoprotein vaccine (124) and an adenovirus vector vaccine (125), have failed in phase III clinical trials. However, preclinical and early clinical data suggested that these vaccine candidates had limited potential to actually prevent infection. Experiments with rhesus macaques using adenovirus vector vaccines did not actually demonstrate protection from infection with SIV (41). Rather, some infected animals who received adenovirus vector vaccines demonstrated lower peak viral load during acute infection and (transiently) lower viral concentrations at set point (41). However, no similar effects were observed in a human trial with an adenovirus vector vaccine. There was a nonsignificant trend toward greater risk for acquisition of HIV-1 in subjects who had received the vaccine (125), especially in subjects with higher concentrations of antibodies directed to the adenovirus vector. Currently, many investigators believe that an HIV-1 vaccine capable of preventing infection is likely to require the generation of neutralizing or other protective antibodies (122, 123), and such a vaccine is unlikely to be available for many years.

Microbicides. Non-vaccine biological interventions designed to prevent infection with HIV-1 after exposure to the virus include topical microbicides. However, multiple first-generation microbicides (nonoxynol-9, cellulose sulphate, and C31G [SAVVY]) have failed to protect in clinical trials (126–129). Indeed, these early microbicides seemed to increase the risk of infection, perhaps through subclinical irritation of the mucosa. Studies to evaluate efficacy of antiretroviral-containing topical microbicides are currently being implemented (109).

Antiviral agents. Perhaps the greatest excitement has been generated by the use of antiretroviral therapy for prevention, as a topical microbicide or as a pre- or postexposure oral prophylaxis to prevent acquisition of HIV-1. The plausibility of this approach is based on the successful use of antiretroviral therapy to prevent mother-to-child transmission (ref. 104 and discussed in detail below) and from a series of animal experiments (130). It has been shown that daily s.c. administration of a combination of the antiviral agents tenofovir and emtricitabine to rhesus macaques using adenovirus vector vaccines did not actually demonstrate partial protection (130). Conversely, tenofovir alone delayed but did not prevent infection with SHIV in these experiments.

To convert these results from animals to human therapy, clinical trials are underway to assure that HIV–1–negative subjects can safely receive antiretroviral agents and determine whether infection with HIV-1 can be averted in high-risk subjects. A summary of ongoing trials is provided in Table 2 (131). One trial demonstrating the safety of tenofovir in women at high risk of becoming infected with HIV-1 in Africa has been completed, but the study did not have the statistical power to assess efficacy of the regimen for preventing infection with HIV-1 (132). However, trials of antiviral drugs for prevention have proven difficult to complete for scientific, logistical, and political reasons (133, 134).

Prevention of mother-to-child transmission during pregnancy, delivery, and breastfeeding (where transmission occurs in 20%–45% of cases without an intervention) has been the most successful HIV-1 prevention intervention to date (11, 135, 136), especially in developed countries, where neonatal infections have been virtually eliminated through routine use of HAART during pregnancy and delivery (135) and where breastfeeding can be safely replaced with infant formula feeding. In addition, use of reliable contraceptives is another effective means to prevent neonatal infections resulting...
from unintended pregnancies in women who are HIV-1–positive (137). Unfortunately, in the developing world, only 11% of pregnant women are estimated to have access to effective HIV-1 prevention services (99) and reliable contraceptives (137).

Attempts to optimize access to antiviral prophylaxis for vertical (mother-to-child) transmission of HIV-1 emphasize the challenges associated with HIV-1 prevention in resource-poor settings. Single-dose nevirapine delivery has been the most important strategy for peripartum prophylaxis because it is simple to administer and has a well-documented benefit (138). However, single-dose nevirapine does not provide as much protection as more complex regimens.

Furthermore, nevirapine resistance has been observed in a substantial number of mothers and babies who have received treatment (139, 140), presumably because the long half-life of the drug gives HIV-1 substantial exposure to a subtherapeutic concentration of the agent and incomplete viral suppression permits selection of drug-resistant HIV-1 variants. Nevirapine resistance that emerges after its use in prophylaxis during pregnancy can compromise the subsequent use of this drug for treatment (141). Strategies are urgently needed to greatly expand access to and maximize the benefits of antiviral prophylaxis for women who are pregnant and neonates (142).

Does antiviral treatment equal prevention?

As indicated above, HIV-1 transmission events transpire from people with acute, established, and late infection. Infection can result from people who do not know their HIV-1 status and from those who persist in risky sexual behavior despite knowledge of underling HIV-1 infection. Massive worldwide testing efforts may help those who test positive to better protect their personal health and avoid spreading the infection. Biological, ecological, and observational data (101) suggest that infected subjects who receive HAART are rendered far less contagious (143). On a population level, the lower the viral burden, the less the anticipated spread of HIV-1 (144). However, it is not clear whether the HIV-1 prevention benefits of HAART for an individual or population can be sustained, or whether they can achieve public health relevance (101). Most recently, a Swiss HIV-1 advisory group has generated guidelines to try to help physicians counsel patients about the potential prevention benefits of antiretroviral therapy, a decision that has generated great controversy (145).

Conclusions

Now, 28 years after HIV/AIDS was first recognized, we have gained remarkable insight into HIV-1 (146). Such information has led to successful treatments that have dramatically reduced HIV-1–related mortality and morbidity and that prevent mother–child transmission of HIV-1. We have recognized the critical importance of marrying treatment and prevention efforts (147, 148), observed the substantial promise of non-vaccine prevention efforts, and, as emphasized in this report, noted subtle but important differences in the spread and biology of HIV-1 in different parts of the world that might help to explain regional differences in HIV-1 prevalence and survival.

Despite these advances, considerable (and in some cases enormous) problems persist. First, access to and utilization of effective interventions remain limited, especially in the developing world. Second, we have identified no means to cure infection with HIV-1, primarily stymied by our inability to eliminate a “latent pool” of cells in which the HIV-1 genome has integrated into the host genome. Last, we have made only little progress in HIV-1 vaccine development given our incomplete understanding of immunity to HIV-1 infection and induction of potent protective immune responses.

This Review was designed to identify the progress made to date and to underscore the challenges that confront us. An entire generation of clinicians, biological scientists, and social scientists have devoted their lives to understanding the biology of HIV transmission and pathogenesis and to developing ways to interrupt the pandemic. The groundwork laid by these individuals sets the stage for the next generation of investigators, who have no choice but to move forward with the greatest urgency.

Acknowledgments

The authors wish to thank Lisa Chensvold and Christopher Hurt (both at the University of North Carolina at Chapel Hill) for their assistance in preparation of this manuscript and Angela Kashuba for assistance in preparation of Figure 4.

Address correspondence to: Myron S. Cohen, University of North Carolina at Chapel Hill, 130 Mason Farm Road, Room 2115, Campus Box 7030, Chapel Hill, North Carolina 27514, USA. Phone: (919) 966-2536; Fax: (919) 966-6714; E-mail: mscohen@med.unc.edu.


