



Published in final edited form as:

Contemp Clin Trials. 2012 November ; 33(6): 1280–1286. doi:10.1016/j.cct.2012.07.007.

Statistical considerations for the HPTN 052 Study to evaluate the effectiveness of early versus delayed antiretroviral strategies to prevent the sexual transmission of HIV-1 in serodiscordant couples

Ying Qing Chen^a, Benoit Masse^b, Lei Wang^a, San-San Ou^a, Xin Li^a, Deborah Donnell^a, Marybeth McCauley^c, Theresa Gamble^d, Heather J Ribauldo^e, Myron S Cohen^f, and Thomas R Fleming^g for the HPTN 052 Study Team

^aStatistical Center for HIV/AIDS Research and Prevention, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N., Seattle, WA 98109, USA

^bCHU-Sainte-Justine, Universite de Montreal, PO Box 6128, Station Centre-ville Montreal, Quebec, H3C 3J7, Canada

^cFamily Health International, 4401 Wilson Blvd, suite 700, Arlington, VA 22203, USA

^dFamily Health International, 2224 East NC Hwy 54, Durham, NC, 27713, USA

^eCenter for Biostatistics in AIDS Research, Department of Biostatistics, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115, USA

^fDivision of Infectious Disease, Department of Medicine, University of North Carolina, 130 Mason Farm Road, 2nd Floor, Chapel Hill, NC 27599, USA

^gDepartment of Biostatistics, University of Washington, F-600, Health Sciences Building, Box 357232, Seattle, WA 98195, USA

Abstract

The HIV Prevention Trial Network (HPTN) 052 Study is a Phase III, two-arm, controlled, open-labeled, randomized clinical trial designed to determine whether early antiretroviral therapy (ART) can prevent the sexual transmission of human immunodeficiency virus type 1 (HIV-1). A total of 1,763 couples in which one partner was HIV-1-positive and the other was HIV-1-negative were enrolled in four continents, nine countries and thirteen study sites. The HIV-1-positive partner was randomly assigned to either of the two arms: “immediate” (early) therapy with ART initiated upon enrollment plus HIV primary care, or “delayed” therapy with HIV primary care but ART initiated when the index case would have two consecutive measurements of a CD4+ cell count within or below the range of 200–250 cells/mm³, or develop an AIDS-defining illness. In this paper, we describe several key statistical considerations for the design of this landmark study. Despite that the observed event rates were lower than expected, which might have compromised the study power, an early release of the trial results in May 2011 showed an overwhelming 96% risk reduction for the immediate therapy in the prevention of genetically linked HIV-1 incident

© 2012 Elsevier Inc. All rights reserved.

Author for Correspondence: Dr. Ying Qing Chen, SCHARP/FHCRC, 1100 Fairview Ave N, M2-C200, Seattle, WA 98109, USA. Tel: (206) 667-7051, fax: (206) 667-4378; yqchen@fhcrc.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

transmissions. Nevertheless, the durability of its long-term effectiveness is yet to be assessed. The HPTN 052 Study is still ongoing and will not complete till 2015.

Keywords

Antiretroviral therapy; Durability; Multinational; Partner; Treatment-as-prevention

1. Introduction

In the absence of antiretroviral therapy (ART), human immunodeficiency virus type 1 (HIV-1) leads to inexorable destruction of critical immune cells (CD4+), opportunistic infections, and death. Since its introduction in the late 1990s, highly active ART has dramatically reduced the morbidity and mortality of HIV-1 infection through sustained reduction in HIV-1 viral replication. Nevertheless, such therapy did not cure HIV infection, and viral resistance was expected to develop in most patients on regimens that were not completely suppressive. Although it was widely recognized that initiation of ART should not be delayed beyond when the CD4+ cell counts fell below 200 cells/mm³, there had been great debate for many years regarding when to start ART to optimize the benefit-to-risk profile of treatment and prevention for both HIV-infected individuals and their sexual partners, respectively. The obvious benefits of ART also were weighed against a global shortage of antiviral agents and treatment infrastructure, cost, short- and long-term side effects and severe consequences of non-adherence.

The HIV Prevention Trial Network (HPTN) 052 Study is a randomized clinical trial in serodiscordant couples to determine whether earlier initiation of ART for HIV-infected (index) participants can reduce the short- and long-term risk of sexual transmission of HIV-1 to their HIV-negative partners and also yield better clinical outcomes in the HIV-infected index participants. The primary objective of the study was to compare the rates of genetically linked HIV-infection among HIV-negative partners of the index cases in the two study arms: immediate ART (initiated immediately upon enrollment, when the CD4+ counts in the index are between 350–550 cells/mm³) and delayed ART (delayed until the participant has two consecutive measurements of a CD4+ cell count below 250 cells/mm³, or develops an AIDS-defining illness); participants in both study arms would receive HIV-1 primary care throughout the study. A key secondary endpoint was the clinical outcome of the HIV-infected index participants, including death, World Health Organization (WHO) Stage 4 events, severe bacterial infections and pulmonary tuberculosis.

On April 28, 2011, an independent data and safety monitoring board (DSMB) of the U.S. National Institutes of Health/National Institute of AIDS and Infectious Disease (NIH/NIAID) Division of AIDS (DAIDS) reviewed the results of an interim analysis of the data collected as of February 21, 2011. Per the DSMB's recommendation, the interim analysis results were released to the trial participants and the general public on May 12, 2011. They showed an overwhelming 96% risk reduction for the immediate therapy in prevention of the genetically linked HIV-1 incident transmissions [[1]]. Since then, all the HIV-positive partners have been provided ART regardless of their CD4 counts.

Despite its early favorable results for the immediate therapy, the HPTN 052 Study is still ongoing to further assess whether or not the early efficacy is durable. It is expected to complete the follow-up of study participants in 2015. In this article, we present several key statistical considerations during the protocol development, which we believe can be helpful in the design of future similar studies.

2. Study design and key statistical considerations

2.1 Study overview

The HPTN 052 Study is a Phase III, two-arm, randomized, controlled, open-labeled and multi-country clinical trial comparing early versus delayed ART strategies for the prevention of HIV transmission in HIV serodiscordant couples and reduction of morbidity and mortality in the HIV-infected index participants. The couples to be enrolled were sexual partners, same or opposite sex, who were married, had been living together, or considered each other a primary partner. They were required to be sexually active. Specifically, couples should have reported to have sex (vaginal or anal) with partner at least 3 times in the last 3 months.

Its study sites include 13 sites in 9 countries: Gaborone, Botswana; Kisumu, Kenya; Lilongwe and Blantyre, Malawi; Johannesburg and Soweto, South Africa; Harare, Zimbabwe; Rio de Janeiro and Porto Alegre, Brazil; Pune and Chennai, India; Chiang Mai, Thailand; and Boston, MA, United States. Although differences might exist among the countries in health delivery practice, all the clinics selected for this study were HPTN Clinical Trial Units that had been prepared to receive same site training.

Between April 2005 and May 2010, 1763 serodiscordant couples were enrolled into the study under its protocol versions 2.0 and 3.0. In order to provide data on long-term ART effectiveness and public health utility, all enrolled couples were to be followed up for at least 5 years, and the HPTN 052 Study is expected to be complete in 2015. All versions of the HPTN 052 Study protocol and its amendments can be found at the HPTN Web site (http://www.hptn.org/research_studies/HPTN052StudyDocuments.asp#Protocol).

2.2 Study endpoints

Genetically linked HIV-1 incident infection occurring in the HIV-negative partners of randomized HIV-infected index participants was the primary prevention endpoint for the study. A complementary analysis would also consider all acquisitions, regardless of their linkage. The effectiveness estimate obtained via this latter analysis would provide a measure of the overall public health effect of ART in the prevention of HIV transmission. Details regarding determination of HIV transmission linkage analyses were provided elsewhere [[2]].

To compare the effectiveness of early versus delayed ART on clinical outcomes in HIV-infected index participants, a primary clinical endpoint was chosen as the earliest event of death, a WHO Stage 4 diagnosis, or a severe bacterial infection or pulmonary tuberculosis. This endpoint reflected the most serious clinical events associated with HIV-1 infection.

The primary prevention endpoint and the primary clinical endpoint are both of time-to-event, which would be analyzed according to a pre-specified time-to-event analysis plan.

Corresponding to the HPTN 052 Study's secondary objectives, several secondary endpoints were specified. Details of these secondary endpoints and how they were measured are provided in Table 1.

2.3 Event rates, sample size and power calculation

The primary objective of obtaining a long-term comparison between treatment arms presented a major challenge to the sample size and power calculation for the HPTN 052 Study. First, the reduction of HIV rates due to ART initiation might not be constant over time, as HIV-infected patients might not adhere to therapy, fail therapy, or develop resistant HIV variants. Second, for index participants in the delayed arm of the trial, the time post-

randomization when ART was initiated would vary since it depends on when the participant had either an AIDS-defining illness or consecutively measured CD4+ cell counts below 250 cells/mm³. Further, the relative reduction of morbidity and mortality in the HIV-infected index partners was also expected to change over time. As a result, any naïve use of conventional methods, for example, the usual Cox proportional hazards model assuming a constant regression parameter, to calculate sample size and power would be problematic.

To account for potentially time-varying ART treatment effectiveness and different ART initiation times in the delayed arm, three main assumptions were assumed: (1) the “baseline” risk of HIV transmission within a couple of the positive partner “not” receiving any ART was expected to decline over time; (2) the ART effectiveness for the positive partner receiving ART might decrease over time during the study follow-up; and (3) the delay time before ART initiation in Arm 2 would have an impact on HIV transmission. To actually calculate the expected cumulative event rates in the two arms and the associated power, we followed adopted a two-step procedure:

- Step 1** The expected differences in cumulative HIV rates at the end of the trial were computed under the assumption that participants in the delayed arm did not initiate ART at any point in time during follow-up. In this step, our calculation would yield an upper limit of power for a given sample size, since that the contrast of ART-initiation would be maximal.
- Step 2** Using the assumption on the delay time before ART initiation, the average effectiveness in Step 1 was re-computed to account for the delayed ART initiation in the delayed arm.

Specifically, we used the following assumptions to facilitate the sample size and power calculation:

- Assumption 1** Table 2 shows the expected annual HIV acquisition rates among the partners of index cases who would receive HIV primary care alone, i.e., no ART initiated at any point in time. In this table, the expected annual HIV acquisition rates were assumed at the aggregate level of all the countries where this trial was conducted, although it was believed that individual countries might observe different rates, for example, due to different distributions of HIV epidemics and different care practices. Based on the expected annual HIV acquisition rates, the 5-year cumulative HIV incidence rate was calculated as 16.6% at the end of the trial.
- Assumption 2** Table 2 also shows the expected ART effectiveness for the immediate arm (Arm 1) compared with HIV primary care alone under the two possible scenarios of decreasing effectiveness over time: high and medium ART effectiveness. This yielded different expected cumulative HIV incidence rates for Arm 1, and different average effectiveness at the end of the trial: 52% for expected high ART effectiveness and 35% for medium ART effectiveness.
- Assumption 3** Table 3 shows the required sample sizes to compare the immediate arm with HIV primary care alone. By assuming a medium effectiveness of ART, the sample size of 1750 would provide 87% power to detect an effectiveness of 35%, leading to an approximate rate reduction of 5.5% in the cumulative HIV rates (from 16.6% to 11.1%). By assuming a high effectiveness of ART, the same sample size of 1750 would provide >99% power to detect an effectiveness of

52%, leading to an approximate rate reduction of 8.3% in the cumulative HIV rates (from 16.6% to 8.3%).

Assumption 4 The index participants randomized to the delayed arm would have ART initiated when their CD4+ cell counts dropped below 250 cells/mm³ during follow-up and/or they had experienced any AIDS-defining illness during follow-up. It was thus expected that such ART initiation during follow-up in the delayed arm would lower the overall risk of acquisition compared with HIV primary care alone. In order to assess the distribution of the time participants would initiate ART during follow-up, we had assumed a uniform distribution of CD4+ cell counts between 350 and 550 in the study population with an average annual loss of 60 cells per year, and 10% annual incidence of AIDS-defining illnesses (independent of CD4+ cell counts). Table 2 shows the expected proportion of participants in the delayed arm who would initiate ART over the follow-up period. From this table, the median time post randomization for the index participants in the delayed arm to have initiated ART was approximately 2.8 years. By the end of the third year of follow-up, we expected that 100% of the index partners in the immediate arm and 56% of those in the delayed arm would have initiated ART.

Assumption 5 We assumed the risk reduction in the delayed arm would take effect at the time of ART initiation. Thus, using assumptions 1 and 4, we could calculate the expected HIV cumulative rates in the delayed arm, as shown in Table 4. For example, the expected cumulative HIV rate for HIV primary care alone reduced to 14.9% from 16.6%, if the ART initiation in the delayed arm was taken into account assuming an expected 25% risk reduction due to ART. It further reduced to 14.2% and 13.2% for the expected risk reductions of 35% and 50%, respectively.

Based on these assumptions, power was calculated (also shown in Table 4) assuming 6.5 years of trial duration, 1.5 years of accrual, and 5% annual loss-to-follow-up per arm. Specifically, for the two scenarios of Assumption 2, in scenario (1), the power was greater than 87% to detect effectiveness > 39%, which amounted to a > 4.9% absolute rate reduction in the cumulative rates (13.2% versus 8.3%). This power was achieved with an upper limit of a 50% risk reduction of acquisition for the partners of index cases who had initiated ART in the delayed arm during follow-up. In scenario (2), the power was 61% to detect a 3.8% absolute rate reduction (14.9% versus 11.1%). If the risk reduction for the partners of those having ART initiated in the delayed arm was more than 25%, the trial would be greatly underpowered. In this case, however, the absolute rate reduction of cumulative HIV rates would be less than 3.1%, which might not be of clinical importance.

In addition to assessing the prevention benefit of reducing new HIV-1 infection, for a compelling result that might change international policy and guidelines, the study should also be powered to determine whether or not earlier initiation of ART provided sufficient reduction in serious clinical events among HIV-infected partners when compared to the delayed ART. Specifically, the study should have a high probability that the upper bound of the 95% confidence interval for the true hazard ratio was <0.8 when the true hazard ratio was in the range of a 40% to 50% reduction. It was determined that the 1750 HIV+ individuals enrolled in the HPTN 052 Study would provide at least 80% power to show that early initiation of ART would provide at least a 20% reduction in hazard of serious clinical events when the true hazard ratio was a 40% reduction, assuming an underlying 5-year event

rate of 18% with delayed ART compared to 9% with immediate ART – an average relative hazard of 0.5.

3. Study monitoring

3.1. General procedure

Both treatment strategies were expected to differentially affect the immunologic and virologic responses throughout follow-up. It was expected that initially the HIV-1 RNA levels would be lower in the immediate arm compared with the delayed arm. Early in the trial, this would lead to a reduced rate of HIV acquisition in the immediate arm. Later in follow-up, these differences might increase, diminish or even be reversed. Therefore, we were expecting short-term differences in effectiveness with a possible reversal in effectiveness in the longer term. Thus, the study data monitoring plan should balance the need to protect trial participants, while enabling the trial to address its primary objective regarding the evaluation of the relative long-term effectiveness of the two intervention strategies.

The HPTN Study Monitoring Committee (SMC) and the NIH/NIAID DSMB have been monitoring the trial by reviewing the study data at least once per year. The SMC reviews mainly focused on the operational characteristics of the trial and the overall trial conduct and performance on data pooled over treatment arms. Since SMC reviews took place prior to DSMB reviews and the SMC review minutes were included in DSMB reports as supplement materials, this review process allowed the DSMB to mostly focus on the review of by-arm efficacy and safety endpoints.

3.2. Monitoring of efficacy and safety endpoints

At study initiation, guidelines were established for monitoring efficacy and safety endpoints at a minimum of three interim and a final analyses to satisfy the ethical need for early study termination if initial results were extreme, while not increasing the chance of false conclusions. Specifically, they should: (1) address the importance that the trial provided persuasive evidence when considering both treatment and prevention issues, (2) adjust for the nature of interim monitoring that involved repeated testing over time, (3) reflect particular caution given the benefit-to-risk profile of an immediate ART strategy relative to a delayed ART strategy could change substantially over time, and (4) be driven by the morbidity and mortality events that had the greatest clinical impact. In addressing these requirements, a composite endpoint for each couple was chosen to be the earliest occurrence of death, a WHO Grade 4 event in the index, such as death or extrapulmonary Tuberculosis, or transmission of HIV to the partner. A time-to-event analysis would be performed for this Mortality/Morbidity (M/M) composite endpoint. In this trial, it was expected that approximately 340 of the 1750 couples would experience an “event” relative to this M/M composite endpoint.

To guide recommendations about trial termination when interim results on the M/M composite endpoint were favorable for the immediate ART strategy, the “upper boundary” to establish superiority for the immediate ART strategy relative to the delayed ART strategy would be based on an application of the O’Brien-Fleming boundary to preserve the (one-sided) 0.025 false positive error rate relative to the hypothesis:

H_0 : the M/M composite endpoint rate for the immediate strategy relative to the delayed strategy is > 0.80

To guide recommendations about trial termination when interim M/M composite endpoint results were unfavorable for the immediate ART strategy, the “lower boundary” to establish

lack of superiority would be based on an application of the O'Brien-Fleming boundary to preserve the (one-sided) 0.025 false negative error rate relative to the hypothesis:

H_1 : the M/M composite endpoint rate for the immediate strategy relative to the delayed strategy is < 0.60

For illustration, Table 5 presents the O'Brien-Fleming boundaries for the relative risk (RR) estimates that would lead to rejection of H_0 or H_1 at analyses performed when one would have observed 25%, 50%, 75% or 100% of the trial's expected total of 340 couples experiencing the M/M composite endpoint.

Observe that, to reach the O'Brien-Fleming boundary when interim results on the M/M composite endpoint were favorable for the immediate ART strategy, the delayed ART group would need to have at least 43 excess M/M composite endpoint events (21 in the immediate ART arm versus 64 in the delayed ART arm) at the 25% information fraction, at least 54 excess events (58 in the immediate ART arm versus 112 in the delayed ART arm) at the 50% information fraction, at least 65 excess events (95 in the immediate ART arm versus 160 in the delayed ART arm) at the 75% information fraction, and at least 74 excess events (133 in the immediate ART arm versus 207 in the delayed ART arm) at the 100% information fraction.

Observe that, to reach the O'Brien-Fleming boundary when interim results on the M/M composite endpoint were unfavorable for the immediate ART strategy, the immediate ART arm would need to have at least 15 excess M/M composite endpoint events (50 in the immediate ART arm versus 35 in the delayed ART arm) at the 25% information fraction, at most 6 fewer events (82 in the immediate ART arm versus 88 on the delayed ART arm) at the 50% information fraction, at most 28 fewer events (113 in the immediate ART arm versus 141 in the delayed ART arm) at the 75% information fraction, and at most 50 fewer events (145 in immediate ART arm versus 195 in the delayed ART arm) at the 100% information fraction.

The Lan-DeMets implementation by DeMets and Lan [[3]] of the O'Brien-Fleming guideline [[4], [5]] was used to provide flexibility in the timing and number of interim analyses that will be performed.

3.3. Monitoring quality of study conduct, operational characteristics and implementation

This trial might also be terminated or modified for poor recruitment, adherence, retention, and/or low HIV acquisition rate. The following measures were intended to serve as guidelines for possibly stopping or modifying the study early, pending a final DSMB recommendation.

Recruitment—The study sites were expected to complete recruitment in 18 months. Given the completion of the run-in period, for the rest of 1668 couples to be recruited in the full study, the study was expected to recruit 60 couples per month for the first 6 months and 110 couples per month thereafter for 12 months. The full study was implemented in a staggered fashion for the study sites due to varying regulatory processes in host countries, which would result in different starting enrollment dates for the sites. Such differences would be taken into account in the recruitment rate calculation. Stopping or modifying the study might be considered if the study team failed to recruit more than 75% of the above targeted rates.

Retention and losses to follow-up—Based on the expected incidence of HIV transmission, the target retention rate for the study was 98% per year, i.e., 2% of loss-to-follow-up per year. Stopping or modifying the study might be considered if the study team failed to retain more than 96% of couples per year, i.e., 4% of loss-to-follow-up per year.

Differential loss-to-follow-up by study arms and sites should be reviewed carefully, since participants might choose to leave the study if treatment appeared to fail and/or if other treatments became available. When the loss-to-follow-up was treated as censored, this type of informative censoring could seriously bias the study's primary analysis if not properly adjusted. This retention guideline might be modified if the baseline incidence was determined to be much lower/higher than expected.

Delay time—The primary objective of this study might not be addressed if the time of ART initiation in the delayed arm was too short. It was expected that ART would be initiated 2 to 3 years (median 2.8 years) after enrollment of participants in the delayed ART arm. Stopping or modifying the study might be considered if the median delay time was less than 1 year. This delay time guideline should be evaluated in light of the safety and efficacy endpoint data collected since the expected differences in HIV transmission relied on assumptions about ART effectiveness over time.

Adherence—The overall ART benefits depended on adherence to the regimens prescribed to suppress viral loads. Direct and indirect measures of adherence would be reviewed:

- Adherence to the study medication, e.g., by pill counts and self-report.
- Measurement of viral load assessed by blood plasma HIV-1 RNA levels. In the first 3 years of follow-up, it was expected that the difference in viral load between treatment arms would be substantial. Stopping or modifying the study should be considered if there was less than a 3-fold difference in viral load between the two arms (favoring the immediate arm) within the first 3 years of follow-up, i.e., less than 0.5 difference in log₁₀ viral load.
- Time of ART initiation since CD4⁺ cell count confirmed to drop below 250 cells/mm³ and/or the occurrence of an AIDS-defining illness in the delayed arm. This was to assess the timely initiation of ART for index cases in the delayed arm once the criteria for initiation had been met. Stopping or modifying the study should be considered if the median time was more than 4 weeks.

Partner's HIV acquisition rate—Pooled (across arms) rate of HIV acquisition would be monitored. Stopping or modifying the study might be considered if the upper boundary of an 80% confidence interval for the pooled rate of HIV acquisition was smaller than the expected rate.

4. Current status of the HPTN 052 Study

As of the writing of this paper, all of the HIV-infected partners of the HPTN 052 Study have received ART. Monitored by the HPTN SMC and NIH/NIAID DSMB, the HPTN 052 Study is actively following up all study participants as planned to be complete in 2015.

5. Discussion

The primary objective of the HPTN 052 Study was to compare, for the prevention of HIV in serodiscordant couples, two ART management strategies: the immediate strategy where ART is initiated immediately following the enrollment of index partners with CD4⁺ counts between 350 and 550, and the delayed strategy where ART is initiated once index partners' CD4⁺ counts drop between 200 and 250 or the occurrence of an AIDS-defining illness during follow-up. To address this objective, the study was designed to be a prospective Phase III two-arm randomized trial with a relatively long period of follow-up. By this design, it was expected that in the short-term the immediate strategy would be better than the delayed regimen in reducing the risk of HIV transmission in a serodiscordant couple,

given the fact that the virus might be significantly suppressed immediately following the ART initiation in the index partner. However, the long term benefit-to-risk ratio of this strategy has yet to be determined. In fact, it was expected that the delayed strategy might lead to improved viral suppression in the longer term. That is, the hazard functions of these two strategies may very well cross over a relatively long time period of follow-up.

Crossing-over of hazard functions is challenging in terms of sample size determination for time-to-event outcomes with a naïve use of the conventional methods based on the log-rank test or the Cox proportional hazards model. For the HPTN 052 Study, we addressed this issue by: (1) comparing the cumulative HIV incidence rates instead of using the usual hazards ratios to estimate treatment effectiveness; (2) assuming non-constant effectiveness of ART in reducing the risk of HIV acquisition and allowing it to vary over time with reduced magnitude to reflect the expectation that ART effectiveness might diminish gradually over time; and (3) including the delay time as the so-called “change point” [[6]] in our calculations for the expected cumulative HIV incidence rates. Since we did not know individual CD4+ counts a-priori, we had calculated a distribution for the ART initiation rate, which was similar to the latent treatment effectiveness lag time discussed in Chen et al (2002) [[7]]. Overall, the above strategies allowed us to compare the risk of HIV acquisition during follow-up while taking into account possible crossover in hazard functions. The strategies were based on a limited number of assumptions resulting from substantial consultation with field experts and literature; however, they did not intend to cover all the possible types of crossover in hazard functions. For example, the timing of ART initiation was assumed to be independent of the duration of partnership. More sophisticated modeling of ART initiation time based on additional characteristics might nevertheless improve the calculation of the delay time distribution.

For the sample size calculations, we had assumed homogenous risk of HIV acquisition, that is, the risk of acquisition for partners of those with CD4+ cell count dropping below 250 cells/mm³ and/or developing AIDS-defining illnesses was similar to the one for partners of those with CD4+ cell count above 250 cells/mm³ with no AIDS-defining illnesses. This assumption could be restrictive. However, if in the absence of ART, the number of infected partners with CD4+ cell count dropping below 250 cells/mm³ and/or experiencing AIDS-defining illnesses was increasing rapidly over time, the overall rates would greatly increase over time. This increase might contradict Assumption 1, which was based on observed data from the literature.

The annual HIV incidence rates of acquisition in Assumption 1 did not necessarily reflect the “true” incidence rates of the HPTN 052 Study population. Higher HIV rates would lead to an increase in power. For example, if the annual rates of Table 2 are 7%, 7%, 5%, 5%, and 2% for Years 1 to 5, respectively, the power under Scenario 1 in Table 7 were all above 90%. For Scenario 2, moderate power between 50% and 75% was achieved if the decrease in risk of acquisition for partners of those on the delayed arm who initiate ART was 25% or 35%.

As of May 12, 2011, when the study results were first disclosed, the actual observed 5-year annual HIV incidence rates were 0.2%, 0.2%, 0.5%, 0% and 0% for the immediate arm, and 2.2%, 2.1%, 2.0%, 7.7% and 0% for the delayed arm. These incidence rates were somewhat below their expected ones, except for the last two years when follow-up data were still limited for the estimates to be reliable. Although they might have compromised the calculated power, the actual treatment difference was so overwhelming that the interim analysis still led to the early disclosure of results.

In addition to the prevention outcomes for the partners, the HPTN 052 Study also gave significant weight to treatment outcomes of the indexes for the immediate and delayed strategies. Finally, extra caution was needed for interim monitoring of the trial since early results would not provide reliable insights on the long-term benefit-to-risk ratio associated with both treatment and prevention outcomes. Hence, our proposed monitoring guidelines based on a composite of treatment and prevention endpoints intended to warrant that the trial would provide persuasive evidence regarding both treatment and prevention issues.

Although the main objective of the HPTN 052 Study was to estimate and compare the effectiveness of the two treatment strategies, treatment effect on behaviors, for instance, was part of the intervention that needed to be included in the assessment of effectiveness. The unblinded nature of its design might allow for the proper and timely clinical management of the index cases. However, statistical bias in response measurements, conscious or unconscious, could occur in particular for the self-reported behavioral and safety data (i.e., symptoms), which would be carefully monitored and investigated.

Nevertheless, the HPTN 052 Study was intended to measure the long-term effectiveness of immediate versus delayed strategies. Although the early release of interim analysis results in favor of the immediate therapy had led to the rest of the ART-naïve index partners in the delayed arm to initiate ART sooner than expected, we observed that there was a sufficient contrast in delay time between the two arms. As a result, the HPTN 052 Study shall continue its course to assess whether or not the early treatment efficacy is durable for the longer term.

Acknowledgments

The HPTN 052 Study is supported by the HIV Prevention Trials Network (HPTN) and by grants (UM1-AI068619 and U01-AI068619; and UM1-AI068617 and U01-AI068617, to the HPTN Statistical and Data Management Center) from the NIH/NIAID.

References

1. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, Hakim JG, Kumwenda J, Grinsztejn B, Pilotto J, Godbole SV, Mehendale S, Chariyalertsak S, Santos BR, Mayer KH, Hoffman IF, Eshleman SH, Piwowar-Manning E, Wang L, Makhema J, Mills LA, de Bruyn G, Sanne I, Eron J, Gallant J, Havlir D, Swindells S, Ribauldo H, Elharrar V, Burns D, Taha TE, Nielsen-Saines K, Celentano D, Essex M, Fleming TR. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med.* 2011; 365:493–505. [PubMed: 21767103]
2. Eshleman S, Hudelson S, Redd A, Wang L, Debes R, Chen YQ, Martens C, Ricklefs S, Selig E, Porcella S, Mushaw S, Ray S, Piwowar-Manning E, McCauley M, Hosseinipour M, Kumwenda J, Hakim J, Chariyalertsak S, de Bruyn G, Grinsztejn B, Kumarasamy N, Makhema J, Mayer K, Pilotto J, Santos B, Quinn T, Cohen M, Hughes J. Analysis of genetic linkage of HIV from couples enrolled in the HIV Prevention Trials Network 052 trial. *J Infect Dis.* 2011; 204:1918–1926. [PubMed: 21990420]
3. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. *StatMed.* 1994; 13:1341–1352.
4. O'Brien PC, Fleming TR. Multiple-testing procedure for clinical trials. *Biometrics.* 1999; 35:549–556.
5. Emerson SS, Fleming TR. Symmetric group sequential test designs. *Biometrics.* 1989; 45:905–23. [PubMed: 2675998]
6. Liang K-Y, Self SG, Liu XH. The Cox proportional hazards model with change point – an epidemiologic application. *Biometrics.* 1990; 3:783–793. [PubMed: 2242414]
7. Chen YQ, Rohde CA, Wang M-C. Additive hazards models with latent treatment effectiveness lag time. *Biometrika.* 2002; 89:914–931.

Table 1

Secondary Endpoints of the HPTN 052 Study

Endpoint	Measurements
Morbidity and mortality of index case	<ul style="list-style-type: none"> • Time from enrollment to time of first occurrence of any HIV/AIDS-related event (see Appendix IV for qualifying events) or death (all causes) • Time from enrollment to death (all causes) • Time from enrollment to time of first occurrence of any HIV/AIDS-related event or WHO disease Stage 2 or 3 (see Appendix IV for qualifying events) • Time from enrollment to time of first occurrence of any other targeted medical condition (see Appendix IV for qualifying events) • Time from enrollment to time of occurrence of any condition outlined in Appendix IV or death
Immunologic response of index case	<ul style="list-style-type: none"> • CD4+ cell count over time • Time from enrollment to immunologic failure. (Immunologic failure is defined as two consecutive measurements of CD4+ cell count at or below the range of 200–250 cells/mm³, or develops an AIDS-defining illness) • Time from initiation of ART to immunologic failure • Time from initiation of secondary regimen to immunologic failure
Virologic response of index case	<ul style="list-style-type: none"> • Blood plasma HIV-1 RNA level over time • Seminal plasma HIV-1 RNA levels over time in males • Cervical vaginal HIV-1 RNA levels over time in females • Time from initiation of the starting regimen to confirmed virologic failure • Time from initiation of secondary regimen to confirmed virologic failure
Initiation of secondary regimen Safety and toxicity of treatment	<ul style="list-style-type: none"> • Time to initiation of secondary regimen (any reason) • Time from enrollment to time of first development and any subsequent occurrence of Grade 3 or 4 ART-related toxicities • Time from enrollment to time of first serious cardiovascular or other metabolic events (Grade 4 and higher) • Time from enrollment to time of first Grade 4 and higher events (any event)
HIV drug resistant virus	<ul style="list-style-type: none"> • Prevalence of drug resistant HIV virus • Proportion of infected partners acquiring a drug resistant HIV virus
Incidence of STDs in index case and partners	<ul style="list-style-type: none"> • Time from enrollment to the time of first development and subsequent development of STDs
Adherence in index case	<ul style="list-style-type: none"> • Adherence to all treatment over time • Adherence to treatment over time following initiation of antiretroviral therapy starting regimen • Adherence to treatment over time following initiation of an antiretroviral therapy secondary treatment regimen
Sexual behavior of index cases on ART, and their partners	<ul style="list-style-type: none"> • Sexual behavior over time following initiation of starting regimen • Sexual behavior over time following initiation of a secondary regimen
Quality-of-Life indicators of index case	<ul style="list-style-type: none"> • Quality-of-Life indicators over time following initiation of starting regimen • Quality-of-Life indicators over time following initiation of a secondary regimen

Table 2

Assumptions on yearly baseline HIV incidence rate^a, yearly ART effectiveness and yearly percentage of ART initiated in the delayed arm

Year(s)	1	2	3	4	5	6	7
Annual HIV incidence rate with primary care alone							
	5%	5%	3%	3%	1%	1%	1%
ART effectiveness							
Assumed High	80%	60%	40%	20%	20%	20%	20%
Assumed Medium	50%	40%	30%	20%	10%	10%	10%
Percentage of ART initiation in the delayed arm							
	10%	27%	56%	80%	100%	100%	100%

^a Annual HIV incidence rates among partners of HIV+ participants receiving primary care alone

Table 3

Sample size for immediate therapy versus primary care alone

Effectiveness	Number of Required Study Couples	
	80% Power	90% Power
50%	640	860
45%	820	1100
40%	1070	1440
35%	1450	1940
25%	3020	4050

Table 4

Power and average effectiveness for HIV incidence endpoint

	Expected cumulative rate in immediate arm	Rate reduction with delayed ART	Expected cumulative rates in delayed arm	Power	Effectiveness
(1) High ART effectiveness	8.3%	25%	14.9%	98%	46%
		35%	14.2%	95%	43%
		50%	13.2%	87%	39%
(2) Medium ART effectiveness	11.1%	25%	14.9%	61%	27%
		35%	14.2%	46%	24%
		50%	13.2%	25%	17%

Table 5

Examples of trial monitoring boundaries

Information Fraction (% of Total Events)	Reject H_0		Nominal one-sided p- values for rejection of H_0	Reject H_1	
	RR	0.80		RR	0.60
25% (85 events)	0.3352	P	0.00001; Z = 4.010		1.4320
50% (170 events)	0.5178	P	0.0023; Z = 2.836		0.9269
75% (255 events)	0.5986	P	0.0103; Z = 2.315		0.8018
100% (340 events)	0.6436	P	0.0225; Z = 2.005		0.7457