31 August 2015

Using observational data to emulate a randomized trial of dynamic treatment switching strategies:

an application to antiretroviral therapy

The ART-CC, the CNICS,

and the HIV-CAUSAL Collaboration*

Correspondence:

Lauren Cain, Department of Epidemiology, Harvard T.H. Chan School of Public Health,

677 Huntington Avenue, Boston MA 02115, USA.

Phone: 617 432 6921; fax: 617 432 1884; e-mail: lcain@hsph.harvard.edu

Word count: 247 (abstract), 3,678 (main text), 1,061 (appendices); 4 tables, 4 appendix tables, 2 figures, 2 appendix figures

Running title: When to switch

* The members of the Writing Committee are listed at the end of the paper.

ABSTRACT

Background: When to switch treatment is an important clinical question, for example, when the current therapy fails or shows suboptimal results. Switching strategies often depend on the evolution of an individual's time-varying covariate(s). These so-called dynamic strategies can be directly compared in randomized trials. For example, consider a trial in which HIV-infected individuals receiving antiretroviral therapy are randomized to switching therapy within 90 days of HIV-1 RNA crossing above a threshold of either 400 copies/mL (tight-control strategy) or 1000 copies/mL (loose-control strategy).

Methods: Here we describe an approach to emulate this trial by applying inverse-probability weighting of a dynamic marginal structural model to observational data from the Antiretroviral Therapy Cohort Collaboration (ART-CC), the Centers for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS), and the HIV-CAUSAL Collaboration.

Results: Of 43,803 individuals who initiated an eligible antiretroviral therapy regimen in 2002 or later, 2,001 met the baseline inclusion criteria for the mortality analysis and 1,641 for the AIDS or death analysis. There were 21 deaths and 33 AIDS or death events in the tight control group, and 28 deaths and 41 AIDS or death events in the loose control group. Compared with tight control, the hazard ratios (95% CI) for loose control were 1.10 (0.73, 1.66) for death and 1.04 (0.86, 1.27) for AIDS or death adjusting for baseline and time-varying variables.

Conclusions: While our effective sample sizes were small and our estimates imprecise, the described methodological approach can serve as an example for future analyses.

KEY MESSAGES

- A hypothetical randomized trial comparing dynamic treatment strategies can be emulated by applying inverse-probability weighting of a dynamic marginal structural model to observational data.
- This approach is facilitated by specifying the protocol of the target trial one would like to emulate in terms of the eligibility criteria, the treatment strategies, the follow-up period, outcomes, causal contrasts of interest, and analysis plan.
- As an example, we apply our approach to compare dynamic switching strategies based on HIV-1 RNA thresholds. In our data, most individuals were doing well on their first-line antiretroviral regimen and no differences between switching at HIV-1 RNA thresholds of 400 and 1,000 copies/mL in preventing death and AIDS-defining illness were detected.

INTRODUCTION

Many clinical decisions involve switching or discontinuing treatment. The most effective switching strategies are dynamic, that is, they involve switching different individuals at different times depending on the evolution of their time-varying covariates. However, very few randomized trials compare two or more dynamic strategies for switching medical treatments. Despite this lack of clinical evidence, many clinical guidelines provide recommendations in the form of dynamic switching strategies.

For example, the guidelines for the management of HIV-infected patients issued by the United States Department of Health and Human Services (DHHS)[1] and the International AIDS Society-USA Panel[2] recommend switching a patient's antiretroviral regimen immediately after a confirmed virologic failure (i.e., 2 consecutive HIV-1 RNA measurements $\geq$ 200 copies/mL), the European AIDS Clinical Society[3] and British HIV Association[4] guidelines recommend switching immediately if HIV-1 RNA > 500-1,000 copies/mL or > 400 copies/mL, respectively (but suggest repeating a viral load measurement if HIV-1 RNA is detectable but below the threshold for switching), and the World Health Organization[5] guidelines recommend waiting to switch until confirmation of HIV-1 RNA > 1,000 copies/mL. This threshold is chosen as it is the lowest level that can be used when measuring viral load from dried blood spots. Tight-control strategies are recommended so as to avoid the use of failing antiretrovirals in the presence of ongoing viral replication which may lead to selection of drug resistant mutations requiring more expensive drugs and limiting future treatment options.[6-9]

Here we review a framework for the comparison of dynamic switching strategies using observational data.[10-14] We begin by describing the protocol of the hypothetical randomized trial we would like to conduct (the target trial). We then review an approach to emulate this target

trial using observational data. To overcome the limitations of standard methods for adjustment for time-varying confounders,[15, 16] we use inverse-probability weighting of a dynamic marginal structural model.[17]

THE PROTOCOL OF THE TARGET TRIAL

The target trial is a hypothetical randomized trial that is specified in order to guide our analysis of observational data. Key components of its design are eligibility criteria, treatment strategies being compared, follow-up period, outcomes, causal contrasts of interest, and analysis plan. These were agreed through discussions between colleagues with clinical and statistical backgrounds, which focused on the hypothetical randomized trial whose results would be most useful to resolve uncertainties in clinical practice. We describe each of these components below.

Eligibility criteria

The trial includes individuals who initiated antiretroviral therapy in 2002 or later, achieved suppression of viral replication (defined as at least one measurement of HIV-1 RNA ≤ 200 copies/mL) within 360 days of initiating treatment, and then experienced confirmed virologic failure (defined as the second of two measurements of HIV-1 RNA > 200 copies/mL 7-180 days apart). At confirmed virologic failure (baseline), individuals are required to be 18 years of age or older and have a CD4 cell count measurement in the previous 90 days. Eligible antiretroviral regimens before first virologic failure are listed in Table 1.

Treatment strategies

Eligible individuals are randomized to either tight- or loose-control strategies at confirmed virologic failure. The tight-control strategy is "switch within 90 days of HIV-1 RNA crossing above 400 copies/mL." The loose-control strategy is "switch within 90 days of HIV-1 RNA crossing above 1000 copies/mL." In both arms, individuals should switch from regimens at baseline to new regimens (as indicated in Table 2) and switches are expected to occur

uniformly[11] during the 90-day grace period. After the switch, individuals may switch to another regimen or discontinue treatment if clinically indicated or recommended by their treating physicians. However, regardless of the treatment received after switching, all individuals should be seen and have their CD4 cell count/HIV-1 RNA measured on average every 12-16 weeks and at least once every 52 weeks. In this target trial, as in all randomized trials, we expect that some individuals will not adhere to their assigned treatment strategy.

Outcomes

The clinical outcomes of interest are all-cause mortality and a combined endpoint of AIDS-defining illness[18] or death.

Follow-up period

Individuals are followed from baseline (randomization) until the outcome, loss to follow-up (52 weeks after the most recent laboratory measurement), or the administrative end of follow-up (3 years after baseline), whichever occurred first.

Causal contrasts of interest

To compare the two switching strategies, we calculate the intention-to-treat effect and the per-protocol effect (i.e., the effect that would have been observed if all participants had switched as indicated in this protocol, regardless of the treatment they received subsequently).

Analysis plan

Intention to treat analysis: We estimate 3-year Kaplan-Meier survival curves by randomization arm. Despite its limitations as an effect measure,[19] we also estimate the mortality hazard ratio via the pooled logistic model logit $\Pr(D_{t+1} = 1|D_t = 0, X) = \beta_{0t} + \beta_1 X$, where D_t is an indicator (1: yes, 0: no) for death in week t, β_{0t} is a time-specific intercept (the baseline hazard, estimated via linear and quadratic terms for t), X is an indicator for randomization arm (1: loose-control, 0: tight-control), and β_1 is the log odds ratio of mortality for loose- versus tight-control. Because mortality is rare in each time interval, the parameter β_1 approximates the log of the intention-to-treat mortality hazard ratio that would have been estimated from a proportional hazards Cox model.[20] In case of a chance imbalance of pre-treatment prognostic factors V between arms, the model would include them as covariates.

Per-protocol analysis: Individuals are censored when they deviate from the switching strategies in this protocol. In particular, individuals are censored at the time they change treatment prematurely (i.e., between baseline and when HIV-1 RNA first crosses above 400 copies/mL for tight control and above 1000 for loose control), change to an ineligible regimen during the 90-day grace period, and at the end of the grace period if the individual has not yet switched to an eligible regimen. Because this censoring may be informative, adjustment for both baseline (pre-randomization) and time-varying (post-randomization) covariates may be necessary.[11]

To estimate the per-protocol average mortality hazard ratio, we fit a weighted model logit $\Pr(D_{t+1} = 1|D_t = 0, C_t = 0, X, V) = \theta_{0t} + \theta_1 x + \theta_2' V$, where C_t is an indicator (1: yes; 0: no) for censoring due to deviating from the assigned switching strategy in week t and V is a vector of the baseline (time-fixed) covariates (sex, age (<35, 35-49, ≥50 years), race (white, black, other or unknown), geographic origin (North America, Western Europe, sub-Saharan Africa, other, or

unknown), mode of acquisition (heterosexual, homosexual/bisexual, injection drug use, other or unknown), CD4 cell count (<200, 200-499, ≥500 cells/mm^3), HIV RNA (≤400, 401-1,000, >1,000 copies/mL), calendar year (2002-2004, 2005-2007, ≥2008), regimen class at initiation (nonnucleoside (NNRTI)-based or non-NNRTI based), and regimen class at baseline (NNRTI-based or non-NNRTI based).

To adjust for time-varying selection bias that is induced by the censoring required for the per-protocol analysis, we use inverse probability weighting to eliminate the dependence between measured prognostic factors and censoring. Informally, an uncensored individual's weight at time *t* is inversely proportional to his/her probability of remaining uncensored through time *t* conditional on having survived to time *t* (D_t), his/her covariate history ($\bar{L}_t$), and his/her switching history ($\bar{A}_{t-1}$). $A_t = 2$ indicates that the individual has switched to an eligible regimen by week *t*, $A_t = 1$ indicates that the individual has changed to an ineligible regimen by week *t*, and $A_t = 0$ indicates that the individual has not changed treatment by week *t*. We weight each individual by the time-varying inverse-probability weight $W_t = \prod_{k=0}^{t} \frac{1}{[f(A_k|\bar{A}_{k-1}, D_k=0, \bar{L}_k)]}$ where $f(A_k|\bar{A}_{k-1}, D_k = 0, \bar{L}_k)$ is the conditional probability mass function $f_{A_k|\bar{A}_{k-1}, D_k=0, \bar{L}_k}(a_k|\bar{a}_{k-1}, d_k = 0, \bar{l}_k)$ with $(a_k|\bar{a}_{k-1}, d_k = 0, \bar{l}_k)$ evaluated at the random argument $(A_k|\bar{A}_{k-1}, D_k = 0, \bar{L}_k)$ and $A_{-1} = 0$.

As previously described,[21] these probabilities are estimated using pooled multinomial logistic models including a time-specific intercept (estimated via linear and quadratic terms for *t*), the baseline covariates previously listed, and the time-varying covariates (CD4 cell count (restricted cubic spline with 5 knots at 10, 200, 350, 500, and 1,000 cells/mm^3), HIV-1 RNA (≤400, 401-1,000, >1,000 copies/mL), AIDS-defining illness (when the outcome was death

alone), and time since last laboratory measurement (<4, 4-7, 8-11, ≥12 weeks)). For an explanation of why the probability of treatment changes can be used to estimate the probability of remaining uncensored, please see Cain et al. 2010.[11]

Under the assumptions that (1) we measured and successfully adjusted for all confounders (i.e., prognostic factors that also predict censoring); (2) there is positivity (i.e., no deterministic treatment assigned given the confounders); and (3) the weight models are not misspecified, the above logistic model estimates the parameter of a dynamic marginal logistic structural model:[17, 22-24] $\Pr(D_{t+1}^{x} = 1 | D_t^{x} = 0, V) = \theta_{0t}^{*} + \theta_1^{*} x + \theta_2^{*\prime} V$ where D_t^{x} is the counterfactual indicator that an individual would have developed the outcome during week *t* under strategy *X*=*x*.

To estimate per-protocol survival curves, we fit a similar model that included a product ("interaction") term between *X* and *f*(*t*) where *f*(*t*) is a flexible function of time (estimated via linear and quadratic terms for *t*). The models' predicted values are then used to estimate the 3-year survival from baseline as previously described.[11, 19] (Nonparametric estimation of survival curves would result in very unstable estimates.)The estimated 3-year survival can be interpreted as the survival that would have been estimated had all individuals switched according to the study protocol (regardless of the treatment they subsequently received).

The same analytic approach is then applied to the combined endpoint of AIDS-defining illness or death. Inverse probability weighting may be used to adjust for potential selection bias due to loss to follow-up[25] in both the intention-to-treat and per-protocol analyses.

EMULATING THE TARGET TRIAL USING OBSERVATIONAL DATA

In the absence of a randomized clinical trial for switching, we emulated one using observational data[22] from the Antiretroviral Therapy Cohort Collaboration (ART-CC), the Centers for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS), and the HIV-CAUSAL Collaboration. These collaborations have been described elsewhere.[26-29] The cohorts that make up these collaborations are listed in Appendix 1. All overlaps between and within collaborations were removed. Each cohort collected data prospectively, including CD4 cell count, HIV-1 RNA (limit of detection ≤ 200 copies/mL), dates of treatment initiation and treatment changes, AIDS-defining illness, and death.

We designed our analysis of the observational data to match the eligibility criteria, the treatment strategies, and the outcomes of the target trial as much as possible.

Eligibility criteria

We applied the same eligibility criteria as in the target trial. Our analysis was restricted to HIV-infected persons who initiated antiretroviral therapy after January 1, 2002 (2004 for CoRIS, 2005 for FHDH and Frankfurt when information on their treatment interruptions became available).

Treatment strategies

We compared the same tight- and loose-control switching strategies as in the target trial. To reduce the influence of data errors, new drug prescriptions of duration 14 days or less were disregarded when determining the existence of switching. Instead, the time was assigned to the nearest regimen of duration longer than 14 days before the short regimen. In sensitivity analyses, point estimates did not vary (data not shown) for durations of 31 and 62 days, when we assigned

the disregarded time to the nearest longer regimen after the short regimen, and when we used an alternative definition of switching (see Table 2).

Outcomes

We considered the same two outcomes as in the target trial: all-cause mortality and a combined endpoint of AIDS-defining illness[18] or death. The date of death was identified using a combination of national and local mortality registries and clinical records as described elsewhere, [28] and AIDS-defining illnesses were ascertained by the treating physicians.

Follow-up period

Follow-up started at baseline and ended at the occurrence of the outcome, loss to follow-up (52 weeks after the most recent laboratory measurement), or the cohort-specific administrative end of follow-up (up to November 2012), whichever occurred first.

Causal contrast of interest

For the reasons explained below, only the per-protocol effect comparing the two switching strategies can be estimated.

Analysis

We used the same pooled logistic model described for the target trial, except that we fitted the model to an expanded data set constructed as follows. Because all individuals had data consistent with both strategies at confirmed virologic failure (baseline), we created an expanded dataset that included two replicates (clones) of each individual, and assigned each replicate to

one of the strategies. We censored replicates if and when their data were no longer consistent with their assigned strategy.[17] In particular, replicates were censored if and when the individual changed treatment too soon (i.e., between baseline and when HIV-1 RNA first crossed above 400 (1000) copies/mL), if and when the individual changed to an ineligible regimen during the 90-day grace period, and at the end of the grace period if the individual had not yet switched to an eligible regimen.

A consequence of using grace periods with cloning and censoring is that an intention-to-treat effect cannot be estimated because each individual is assigned to all strategies at baseline. Therefore, a contrast based on baseline assignment (i.e., an intention-to-treat analysis) will compare groups with essentially identical outcomes. Analyses with a grace period at baseline are geared towards estimating a per-protocol effect of a target trial.

The inverse-probability weights were the same as for the target trial except that we added a numerator[11] to emulate uniform switching during the grace period. This numerator equals $\frac{1}{m+1-j}$ when $j = m$ and when $0 \leq j < m$ if the individual initiates and $1 - \frac{1}{m+1-j}$ when $0 \leq j < m$ if the individual does not initiate where m is the length of the grace period in weeks and j is the position in the grace period such that $j = 0$ is the beginning of the grace period and $j = m$ is the end of the grace period. The weights were truncated at the 99th percentile;[30] however, truncation had little effect on the estimates (data not shown).

The emulation of the design and analysis of the alternative trial in which we would not require confirmation of virologic failure was identical, except that baseline was the time of first virologic failure. The inclusion of inverse-probability weights to adjust for censoring at 52 weeks without a laboratory measurement in addition to the previously described weights had little effect on our estimates (results not shown).

All 95% CIs were estimated via a nonparametric bootstrap with 500 samples. All analyses were conducted with SAS 9.4 (SAS Institute, Cary, North Carolina, USA).

RESULTS

Of 43,803 potentially eligible individuals, 2,001 met the baseline inclusion criteria for the mortality analysis and 1,641 for the AIDS or death analysis. The most common reason for being excluded was never experiencing virologic failure after achieving virologic suppression. A flowchart of patients for the mortality analysis is provided in Figure 1.

Table 3 shows the baseline characteristics of the study population for the mortality analysis. Of the 4,002 replicates in the expanded dataset for the mortality analysis, 74% of the tight control group and 68% of the loose control group were censored during follow-up. In the tight control group, 11% were censored for changing treatment prematurely, 14% were censored for changing to an ineligible regimen during the grace period, and 75% were censored for not having switched to an eligible regimen by the end of the grace period. In the loose control group, 23% were censored for changing treatment prematurely, 14% were censored for changing to an ineligible regimen during the grace period, and 63% were censored for not having switched to an eligible regimen by the end of the grace period. Among the uncensored, the median (IQR) follow-up time was 89 (38, 168) weeks for the tight control group (1,673 person-years) and 82 (40, 166) weeks for the loose control group (2,009 person-years). The numbers were similar in the AIDS or death analysis.

There were 21 deaths and 33 AIDS or death events in the tight control group, and 28 deaths and 41 AIDS or death events in the loose control group (Table 4; see Appendix 2 for details). Among those who died, the median (IQR) time to death was 31 (11, 52) weeks for the tight control group and 42 (14, 113) weeks for the loose control group. Among those who developed AIDS or died, the median (IQR) time to AIDS or death was 11 (2, 29) weeks for the tight control group and 15 (7, 60) weeks for the loose control group. Compared with tight

control, the fully-adjusted hazard ratios (95% CI) for loose control were 1.10 (0.73, 1.66) for death and 1.04 (0.86, 1.27) for AIDS or death. Adjustment for either baseline or time-varying variables did not materially change the hazard ratio estimates (Table 5). The estimated inverse probability weights for the mortality analysis had mean 3.1 (interquartile range 1.2 - 3.2, 99th percentile 15.5). The estimated inverse probability weights for the AIDS or death analysis had mean 3.1 (interquartile range 1.1 - 3.4, 99th percentile 17.2). The main predictors of switching to an eligible regimen and changing to an ineligible regimen were time-varying HIV-1 RNA and time since last laboratory measurement (see Appendix Table 1).

Figure 2 plots the estimated 3-year survival and 3-year AIDS-free survival. The survival at 3 years was 95.7% (93.4%, 98.1%) for tight control and 95.2% (92.8%, 97.6%) for loose control. The 3-year survival difference was -0.5% (-2.3%, 1.2%). The AIDS-free survival proportion was 93.3% (90.5%, 96.1%) for tight control and 92.8% (89.7%, 95.9%) for loose control. The 3-year AIDS-free survival difference was -0.5% (-1.9%, 0.8%).

As a sensitivity analysis, we also considered an alternative trial in which we did not require confirmation of virologic failure. In this case, baseline becomes the time of first virologic failure (defined as one measurement of HIV-1 RNA > 200 copies/mL) following virologic suppression. Estimated hazard ratios using this definition of baseline were similar (see Appendix 3 for details).

DISCUSSION

We have described how to use observational data to emulate a hypothetical randomized trial comparing different treatment switching strategies. As an illustration, we applied the method to the question of when to switch from a first-line antiretroviral regimen to a new regimen following virologic failure.[31]

Our results suggest that there is little difference between switching within 90 days of HIV-1 RNA crossing above a threshold of either 400 copies/mL or 1000 copies/mL in terms of preventing short-term death and AIDS-defining illness. However, even after pooling data from three large consortia of HIV cohorts, our effective sample size was small and the effect estimates imprecisely estimated. This was due, in large part, to the strict eligibility criteria of our target trial, which were defined by a panel of clinicians on the basis of the treatment guidelines. Of the 43,803 potentially eligible individuals, 95% were excluded because they did not meet the baseline inclusion criteria. Most individuals excluded were doing well on their first-line antiretroviral regimen and did not experience virologic failure. Had we been able to observe individuals for longer periods of time, more of them would likely have experienced virologic failure and could have been included in our analyses.

Most individuals in our analysis contributed to both arms of the target trial because one cannot generally observe the exact moment at which these HIV-1 RNA thresholds were crossed. As a result, 59% of individuals crossed both thresholds at baseline (those with baseline HIV-RNA $\leq$ 400 copies/mL had the potential to cross both thresholds simultaneously later in their follow-up). In the main analysis, 20 of 29 individuals who died and 31 of 43 individuals who developed AIDS or died contributed events to both groups (see Appendix 2 for details). Similar

difficulties have been encountered when trying to emulate target trials that compare two dynamic strategies in cancer patients.[31]

The validity of our methodology relies on two key assumptions in addition to positivity. First, we assume there is no unmeasured confounding given the measured covariates, i.e., that all joint predictors of switching and the outcome were included in the estimation of the inverse probability weights. The assumption might not hold, even approximately, if for example prior adherence to treatment and antiretroviral drug resistance remained important predictors of treatment switching and the outcome even after adjustment for the measured covariates (some of which may be viewed as proxies for adherence and resistance). To further protect our estimates from unmeasured confounding, we defined the dynamic treatment strategies in terms of initial switching regardless of subsequent adherence to treatment. Defining the strategies this way makes it unnecessary to adjust for joint determinants of future switching, and is perhaps more clinically meaningful, as at the time of deciding whether or not to switch, future adherence is unknown.

Second, we assumed a correct specification of the model for switching as a function of the measured confounders. To reduce bias due to model misspecification that results in apparent outliers, we truncated the estimated weights at the 99th percentile of the distribution of the estimated weights.[30]

Our analyses only focused on the decision to switch regimens after treatment failure, but in practice switching may occur for other reasons, including regimen simplification, toxicity management, and avoidance of teratogenic effects during pregnancy. While the dates of pregnancies were not available for the majority of individuals in this analysis, we restricted the analysis to those who became virologically suppressed and therefore were more likely to adhere

and less likely to experience treatment-related toxicities (more common in the early stages of therapy).

We defined our treatment strategies for switching based on HIV-1 RNA viral load only. The majority of clinical guidelines[1-4] also recommend investigating the reasons for failure, addressing any adherence issues, and performing resistance testing while the individual is on the failing regimen before switching. While data on adherence and the results of resistance testing were not available for the majority of individuals in this analysis, we hope to be able to incorporate these data in the future. These considerations may suggest that even with reasonable eligibility criteria and minimal unmeasured confounding, our target trial was of limited clinical relevance in the populations and periods during which the observational data for our study were collected.

In summary, we described an approach to compare dynamic strategies of treatment switching via censoring and inverse probability weighting. We expect that the methodological approach described here for the comparison of dynamic switching strategies using observational data will serve as an example for future analyses. Future applications may consider switching strategies for which more HIV-infected individuals are eligible and the use of alternative methods for comparing dynamic strategies of treatment switching, including the parametric g-formula, that may result in more precise estimates at the expense of additional modeling assumptions.[16, 32-34]

APPENDIX 1

The Antiretroviral Therapy Cohort Collaboration (ART-CC) includes 20 prospective cohort studies from 38 countries: FHDH (French Hospital Database on HIV), ICONA (Italian Cohort of Antiretroviral-Naïve Patients), SHCS (Swiss HIV Cohort Study), ATHENA (AIDS Therapy Evaluation project Netherlands), The Multicenter Study Group on EuroSIDA, Vanderbilt HIV Cohort (USA), Frankfurt HIV Cohort (Germany), Aquitaine Cohort (France), British Columbia Center for Excellence in HIV/AIDS (Canada), Royal Free Hospital Cohort (UK), Southern Alberta Clinic (Canada), Köln/Bonn cohort (Germany), PISCIS (Proyecto para la Informatización del Seguimiento Clínico-epidemiológico de la Infección por HIV y SIDA, Spain), 1917 Clinic Cohort (University of Alabama, Birmingham, USA), University of Washington HIV Cohort (Seattle, WA, USA), VACS (Veterans Aging Cohort Study, USA), HAVACS (HIV Atlanta Veterans Affairs Cohort Study, USA), CoRIS (Cohorte de la Red de Investigación en SIDA, Spain), VACH (Spain), and OEHIVKOS (Österreichische HIV-Kohortenstudie, Austria).

The Centers for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS) includes 8 clinical cohort studies from the United States. Two of these cohorts also appear in ART-CC: the 1917 Clinic Cohort and the University of Washington HIV Cohort. The others do not: Case Western Reserve University, University of California, San Francisco, the University of California, San Diego, Fenway Community Health Center of Harvard University, University of North Carolina, and Johns Hopkins University.

The HIV-CAUSAL Collaboration includes 13 prospective cohorts from 6 European countries and the United States. Seven of these cohorts also appear in ART-CC: FHDH, SHCS, ATHENA, Aquitaine, PISCIS, VACS, CoRIS. The other six cohorts do not: UK CHIC (United

Kingdom Collaborative HIV Cohort), AMACS (Athens Multicenter AIDS Cohort Study, Greece), UK Register of HIV Seroconverters (UK), ANRS PRIMO and ANRS SEROCO (Agence Nationale de Recherches sur le SIDA, France), and GEMES (Grupo Español Multicéntrico para el Estudio de Seroconvertores, Spain).

Appendix Figure 1 lists the cohorts from each collaboration and shows any overlaps. Appendix Table 2 provides the distribution by cohort of the 43,803 individuals at initiation and 2,001 individuals at baseline in the mortality analysis.

APPENDIX 2

The expanded dataset included two replicates (clones) of each individual because all individuals had data consistent with both strategies at confirmed virologic failure (baseline). Each replicate was assigned to one of the strategies, either tight or loose control. As a result of the expansion, it is possible for an individual to contribute an outcome to one or both strategies. In the mortality analysis, 29 individuals contributed 49 deaths. Of the 29 individuals, 1 contributed a death to the tight control group only, 8 contributed deaths to the loose control group only, and 20 contributed deaths to both groups. In the AIDS or death analysis, 43 individuals contributed 74 AIDS or death events. Of the 43 individuals, 2 contributed an event to the tight control group, 10 contributed events to the loose group only, and 31 contributed events to both groups.

Appendix Table 3 displays key dates and HIV-1 RNAs for the 29 individuals who contributed deaths to the mortality analysis. In addition to the death date, the baseline date, the date at which the threshold for tight control was crossed (if any), and the date at which the threshold for loose control was crossed (if any) are given along with the HIV-RNA at these time points. If the individual switched to an eligible new regimen within the grace period, a switch date is also provided.

Individual 1's death contributed to the tight control group only. Individual 1crossed the tight control threshold at baseline and then switched within the grace period. He was censored from the loose control group at the time of his switch. The deaths of individuals 2-9 counted towards the loose control group only. Individuals 2-4 crossed the tight control threshold after baseline. They were censored from the tight control group 90 days later. They then died without crossing the loose control threshold or switching. Individuals 5-8 crossed the tight control

threshold at baseline. They were censored from the tight control group 90 days after baseline. They then died without crossing the loose control threshold or switching. Individual 9 crossed the tight control threshold at baseline. He was censored from the tight control group 90 days after baseline. He later crossed the loose control threshold, switched within the grace period, and died.

The deaths of individuals 9-22 counted towards both the loose and tight control groups. Individuals 10-15 died without crossing either threshold or switching. Individuals 16-20 crossed the thresholds for both tight and loose control at baseline and later died during the grace period without switching. Individuals 21-28 crossed both thresholds at baseline, switched within the grace period, and died. Individual 29 crossed both thresholds after baseline, switched within the grace period, and died.

APPENDIX 3

When not requiring confirmation of virologic failure, there were 6,320 individuals who met the baseline inclusion criteria for the mortality analysis and 5,310 for the AIDS or death analysis. Of the 12,640 replicates in the mortality analysis, 76% of the tight control group and 68% of the loose control group were censored. Among the uncensored, the median (IQR) follow-up time was 82 (35, 157) weeks for the tight control group and 80 (38, 154) weeks for the loose control group.

There were 50 deaths and 68 AIDS or death events in the tight control group, and 63 deaths and 83 AIDS or death events in the loose control group (Appendix Table 4). Among those who died, the median (IQR) time to death was 34 (8, 80) weeks for the tight control group and 37 (9, 87) weeks for the loose control group. Among those who developed AIDS or died, the median (IQR) time to AIDS or death was 9 (3, 46) weeks for the tight control group and 19 (5, 57) weeks for the loose control group. Compared with tight control, the fully-adjusted hazard ratios (95% CI) for loose control were 0.86 (0.55, 1.33) for death and 1.06 (0.79, 1.43) for AIDS or death.

Appendix Figure 2 plots the estimated 3-year survival and 3-year AIDS-free survival when we did not require confirmation of virologic failure. The survival at 3 years was 96.5% (94.5%, 98.4%) for tight control and 97.0% (95.7%, 98.3%) for loose control. The 3-year survival difference was 0.5% (-1.0%, 2.0%). The AIDS-free survival proportion was 94.9% (92.8%, 97.0%) for tight control and 94.6% (92.6%, 96.6%) for loose control. The 3-year AIDS-free survival difference was -0.3% (-1.9%, 1.4%).

Writing Committee:

Lauren E. Cain, Michael S. Saag, Maya Petersen, Margaret T. May, Suzanne M. Ingle, Roger Logan, James M. Robins, Sophie Abgrall, Bryan E. Shepherd, Steven G. Deeks, M. John Gill, Giota Touloumi, Georgia Vourli, François Dabis, Marie-Anne Vandenhende, Peter Reiss, Ard van Sighem, Hasina Samji, Robert Hogg, Jan Rybniker, Caroline A. Sabin, Sophie Jose, Julia del Amo, Santiago Moreno, Benigno Rodriguez, Alessandro Cozzi-Lepri, Stephen L. Boswell, Christoph Stephan, Santi Pérez-Hoyos, Inma Jarrin, Jodie L. Guest, Antonella D'Arminio Monforte, Andrea Antinori, Richard Moore, Colin N.J. Campbell, Jordi Casabona, Laurence Meyer, Rémonie Seng, Andrew Phillips, Heiner C. Bucher, Matthias Egger, Michael Mugavero, Richard Haubrich, Elvin Geng, Ashley Olson, Joseph J. Eron, Sonia Napravnik, Mari M. Kitahata, Stephen E. Van Rompaey, Ramón Teira, Amy C. Justice, Janet Tate, Dominique Costagliola, Jonathan A. Sterne, Miguel A. Hernán

Funding:

The ART-CC is supported by the UK Medical Research Council [grant numbers G0700820, MR/J002380/1] and the Department for International Development (DFID). The CNICS is support by the National Institutes of Health [grant number R24 AI067039] that was made possible by the National Institute of Allergy and Infectious Diseases (NIAID) and the National Heart, Lung and Blood Institute (NHLBI). The HIV-CAUSAL Collaboration is supported by the National Institutes of Health [grant number R01-AI073127]. Jonathan Sterne is funded by National Institute for Health Research Senior Investigator award NF-SI-0611-10168.

Conflicts of Interest:

Andrea Antinori currently receives honoraria for consultancy agreement by Gilead Sciences, Bristol Myers Squibb, Merck, Janssen-Cilag, ViiV Healthcare, AbbVie. Her institution receives research grants by Gilead Sciences, Bristol Myers Squibb, Janssen-Cilag. All these conflicts are outside of the submitted work.

Heiner Bucher has received honorarium from BMS and Gilead Sciences in the past 6 months. His institution has received grants from BMS And Gilead Sciences. His institution has received funds for travel reimbursement from Gilead Sciences and ViiV Healthcare.

Joseph J. Eron is a consultant to Merck, BMS, ViiV Healthcare, Janssen and Gilead Sciences. His university receives grant support in his name from Janssen and ViiV Healthcare.

Richard Haubrich is an employee of Gilead Sciences.

Michael Mugavero is a consultant with Janssen Therapeutics, Georgetown University / amfAR, NASTAD, Bristol-Myers Squibb, and Gilead Sciences, and has received fees for his services as a consultant.

Peter Reiss through his institution has received independent scientific grant support from Gilead Sciences, Janssen Pharmaceutical Inc, Merck & Co, Bristol-Myers Sqiubb and ViiV Healthcare; he has served on scientific advisory board for Gilead Sciences; he serves on data safety monitoring committee for Janssen Pharmaceutical Inc; chaired a scientific symposium by ViiV Healthcare, for which his institution has received remuneration.

Michael Saag has received research grant support paid to his institution from Merck, Gilead, BMS, ViiV, and AbbVie.

Caroline Sabin is a member of the speakers' bureau for Gilead Sciences. She provides educational training materials for Gilead Sciences, ViiV Healthcare and Janssen-Cilag. She is a member of data safety and Advisory Boards for Janssen-Cilag and ViiV Healthcare and has given talks for Bristol Myers Squibb.

Ramon Teira has occasionally served as a speaker in symposia or other events organized by: Janssen Cilag, Abbvie, ViiV Healthcare and Gilead Sciences.

Table 1: Eligible initial regimens

Regimen Classification	Eligible initial regimens*
PI + ≥2 NRTI	all regimens where the PI is either fosamprenavir (FAPV) or atazanavir (ATV) except those containing the NRTI tenofovir (TNV) or an excluded drug†
bPI + ≥2 NRTI	all regimens except those containing an excluded drug†
NNRTI + ≥2 NRTI	all regimens except those containing an excluded drug†
<6 drugs including FI/INSTI (+ entry inhibitors)	all drug regimens with ≥3 drugs except those containing an excluded drug†

Abbreviations: PI, protease inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; bPI, boosted protease inhibitor; NNRTI, non-nucleoside reverse transcriptase inhibitor; FI, fusion inhibitor; INSTI, integrase strand transfer inhibitor

* Eligible regimens were determined by a panel of clinicians on the basis of treatment guidelines.

† The following drugs are excluded from initial regimens: enfuvirtide (ENF), zalcitabine (DDC), tipranavir (TPV), alovudine (ALO), capravirine (CPV), DPC 083 (DPC083), delavirdine (DLV), emivirine (EMV), lodenosine (DDA or LDN), loviride (LOV), mozenavir (MOZ), vicriviroc (VIC), and any unspecified drugs (ART, PI, NNRTI, NRTI).

Table 2: Changes from initial regimens (columns) to new regimens (rows) that are considered switches*

Regimen Classification	Switch from (PI + ≥2 NRTI)?	Switch from (bPI + ≥2 NRTI)?	Switch from (NNRTI + ≥2 NRTI)?	Switch from (<6 drugs including FI/INSTI + (entry inhibitors))?
PI + ≥2 NRTI	no†	no	yes	yes
bPI + ≥2 NRTI	yes	yes if PI changes†	yes	yes
NNRTI + ≥2 NRTI	yes	yes	yes if NNRTI to etravirine (ETV) †	yes
bPI + PI/NNRTI (+ other)	yes	yes	yes if NNRTI to etravirine (ETV) †	yes
<6 drugs including FI/INSTI (+ entry inhibitors)	yes	yes	yes	yes if FI/II/entry inhibitor changes or addition of a FI/II/entry inhibitor†

If the cell reads "no" this type of change is never considered a switch. Changes to regimen classifications other than those in the table are never switches. If the cell reads "yes" this type of change is always considered a switch. If the cell reads "yes if…" the condition(s) listed must be met for the change to be considered a switch. Other aspects of the regimen may also change or stay the same.

Abbreviations: PI, protease inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; bPI, boosted protease inhibitor; NNRTI, nonnucleoside reverse transcriptase inhibitor; FI, fusion inhibitor; INSTI, integrase strand transfer inhibitor

* Eligible regimens were determined by a panel of clinicians on the basis of treatment guidelines.

† Our primary definition of switching above does not include NRTI-only changes. An alternative definition includes some NRTI-only changes (i.e., any NRTI to tenofovir (TNV) and tenofovir (TNV) to zidovudine (AZT)). According to this alternative definition, a change where the regimen classification does not change is considered a switch if any part of the regimen changes (according to the conditions above).

Finally, individuals must change to regimens that do not include any of the following drugs: zalcitabine (DDC), alovudine (ALO), capravirine (CPV), DPC 083 (DPC083), delavirdine (DLV), emivirine (EMV), lodenosine (DDA or LDN), loviride (LOV), mozenavir (MOZ), vicriviroc (VIC), or any unspecified drugs (ART, PI, NNRTI, NRTI).

Table 3: Characteristics of 2,001 HIV-infected individuals in the mortality analysis at baseline in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

Characteristic		No. of individuals (%)			
		Western Europe 1,503		North America 498	
Sex	Male	959	(63.8)	418	(83.9)
	Female	544	(36.2)	80	(16.1)
Age, years	< 35	540	(35.9)	92	(18.5)
	35 – 50	774	(51.5)	278	(55.8)
	> 50	189	(12.6)	128	(25.7)
Geographic Origin	North America	0	(0)	498	(100.0)
	Western Europe	587	(39.1)	0	(0)
	Sub-Saharan Africa	516	(34.3)	0	(0)
	Other/Unknown	400	(26.6)	0	(0)
Race	White	509	(33.9)	176	(35.3)
	Black	398	(26.5)	215	(43.2)
	Other/Unknown	596	(39.7)	107	(21.5)
Acquisition group	Heterosexual	836	(55.6)	102	(20.5)
	Homosexual	482	(32.1)	158	(31.7)
	Injection drug use	97	(6.5)	81	(16.3)
	Other/Unknown*	88	(5.9)	157	(31.5)
CD4 cell count, per mm^3	< 200	387	(25.7)	150	(30.1)
	200 – 499	801	(53.3)	250	(50.2)
	≥ 500	315	(21.0)	98	(19.7)
HIV-1 RNA, copies/mL	≤ 400	317	(21.1)	147	(29.5)
	401 - 1000	270	(18.0)	95	(19.1)
	>1,000	916	(60.9)	256	(51.4)
Calendar year	2002 – 2004	256	(17.0)	157	(31.5)
	2005 – 2007	714	(47.5)	226	(45.4)
	≥ 2008	533	(35.5)	115	(23.1)

Regimen Class at Initiation	PI + ≥ 2 NRTI	8	(0.5)	9	(1.8)
	bPI + ≥ 2 NRTI	703	(46.8)	219	(44.0)
	NNRTI + ≥ 2 NRTI	785	(52.2)	268	(53.8)
	<6 drugs including FI/INSTI (+ entry inhibitors)	7	(0.5)	2	(0.4)
Regimen Class at Baseline	PI + ≥ 2 NRTI	15	(1.0)	10	(2.0)
	bPI + ≥ 2 NRTI	732	(48,7)	229	(46.0)
	NNRTI + ≥ 2 NRTI	747	(49.7)	257	(51.6)
	<6 drugs including FI/INSTI (+ entry inhibitors)	9	(0.6)	2	(0.4)

Abbreviations: HIV, human immunodeficiency virus; PI, protease inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; bPI, boosted protease inhibitor; NNRTI, nonnucleoside reverse transcriptase inhibitor; FI, fusion inhibitor; INSTI, integrase strand transfer inhibitor

* Other/Unknown acquisition group included all VACS-VC participants includes all VACS participants.

Table 4: Hazard ratios of clinical outcomes under tight and loose control switching strategies in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

Outcome	Strategy (HIV-1 RNA threshold in copies/mL)	No. of outcomes (overlap with tight)	Median (IQR) time to event in weeks	Hazard Ratio, 95% confidence interval					
				Unadjusted		Baseline-adjusted		Baseline and time-varying adjusted* †	
Death	Tight (400)	21	31 (11, 52)	1 (ref.)		1 (ref.)		1 (ref.)	
	Loose (1,000)	28 (20)	42 (14, 113)	1.11	0.88, 1.41	1.13	0.93, 1.38	1.10	0.73, 1.66
AIDS or death	Tight (400)	33	11 (2, 29)	1 (ref.)		1 (ref.)		1 (ref.)	
	Loose (1,000)	41 (31)	15 (7, 60)	1.08	0.90, 1.28	1.05	0.90, 1.23	1.04	0.86, 1.27

Abbreviations: HIV, human immunodeficiency virus; IQR, interquartile range; ref, reference

* Adjusted for the baseline covariates (sex, age, race, geographic origin, mode of acquisition, CD4 cell count, HIV RNA, calendar year, regimen class at initiation, and regimen class at baseline) and time-varying covariates (CD4 cell count, HIV RNA, AIDS-defining illness, and time since last laboratory measurement).

† Time-varying adjustment carried out by inverse probability weighting with weights truncated at the 99th percentile.

Appendix Table 1: Hazard ratios of changing to an ineligible regimen and switching to an eligible regimen in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

	Characteristic		Hazard Ratio, 95% confidence interval			
			Change to an ineligible regimen		Switch to an eligible regimen	
	Baseline hazard	per 13 weeks (linear)	0.88	0.84, 0.92	0.91	0.86, 0.96
		per 13 weeks (quadratic)	1.00	1.00, 1.01	1.00	1.00, 1.01
Baseline Covariates	Sex	Male	1.10	0.86, 1.39	0.96	0.77, 1.18
		Female	(ref.)		(ref.)	
	Age, years	< 35	1.28	0.96, 1.71	0.98	0.75, 1.28
		35 – 50	1.14	0.88, 1.49	1.14	0.90, 1.44
		> 50	(ref.)		(ref.)	
	Residence	North America	1.18	0.85, 1.62	0.71	0.52, 0.97
		Western Europe	1.30	0.93, 1.81	1.29	0.95, 1.75
		Sub-Saharan Africa	1.06	0.78, 1.45	1.27	0.98, 1.66
		Other/Unknown	(ref.)		(ref.)	
	Race	White	1.00	0.76, 1.32	0.67	0.51, 0.87
		Black	1.23	0.95, 1.59	0.91	0.72, 1.15
		Other/Unknown	(ref.)		(ref.)	
	Acquisition group	Heterosexual	1.52	1.08, 2.13	1.14	0.84, 1.55

		Homosexual	1.65	1.17, 2.32	1.37	1.00, 1.88
		Injection drug use	1.45	0.98, 2.14	1.06	0.73, 1.55
		Other/Unknown*	(ref.)		(ref.)	
	CD4 cell count, per mm^3	< 200	1.00	0.66, 1.49	0.92	0.62, 1.37
		200 – 499	1.15	0.85, 1.57	1.24	0.92, 1.67
		≥ 500	(ref.)		(ref.)	
	HIV-1 RNA, copies/mL	≤ 400	1.00	0.77, 1.29	1.11	0.86, 1.43
		401 - 1000	1.09	0.85, 1.4	0.91	0.70, 1.17
		>1,000	(ref.)		(ref.)	
	Calendar year	2002 – 2004	1.98	1.52, 2.58	0.81	0.64, 1.03
		2005 – 2007	1.46	1.15, 1.85	0.90	0.74, 1.09
		≥ 2008	(ref.)		(ref.)	
	Regimen Class at Initiation	non-NNRTI based	1.12	0.81, 1.55	1.12	0.85, 1.48
		NNRTI-based	(ref.)		(ref.)	
	Regimen Class at Baseline	non-NNRTI based	0.79	0.57, 1.09	0.59	0.45, 0.78
		NNRTI-based	(ref.)		(ref.)	
Time-varying covariates	CD4 cell count, per 10 mm^3	Restricted cubic spline: 5 knots at 10, 200, 350, 500, 1,000 mm^3	1.04	1.00, 1.08	0.96	0.93, 0.99
			0.64	0.46, 0.88	1.11	0.84, 1.47
			3.82	1.35, 10.80	0.86	0.34, 2.19
			0.31	0.10, 0.97	0.98	0.33, 2.92

HIV-1 RNA, copies/mL	≤ 400	0.43	0.34, 0.54	0.25	0.20, 0.32
	401 - 1000	0.48	0.33, 0.69	0.63	0.46, 0.86
	>1,000	(ref.)		(ref.)	
Time since last laboratory measurement (weeks)	<4	4.52	3.35, 6.09	12.17	8.08, 18.32
	4 - 7	1.91	1.36, 2.68	4.89	3.17, 7.55
	8 - 11	1.01	0.67, 1.55	2.84	1.74, 4.62
	≥ 12	(ref.)		(ref.)	
AIDS-defining illness	No	0.96	0.86, 1.07	0.90	0.73, 1.11
	Yes	(ref.)		(ref.)	

Abbreviations: HIV, human immunodeficiency virus; NNRTI, nonnucleoside reverse transcriptase inhibitor; ref, reference
* Other/Unknown acquisition group included all VACS-VC participants includes all VACS participants.

Appendix Table 2: Distribution by cohort of the 43,803 individuals at initiation and 2,001 individuals at baseline in the mortality analysis in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

Cohort	No of individuals (%)			
	Eligible at initiation		Eligible at baseline	
	43,803		2,001	
AMACS	705	(1.6)	17	(0.9)
ATHENA	3314	(7.6)	97	(4.9)
Alberta	360	(0.8)	7	(0.4)
Aquitaine	501	(1.1)	13	(0.7)
BCCfE	1146	(2.6)	34	(1.7)
CBC	521	(1.2)	25	(1.3)
CHIC	13693	(31.3)	699	(34.9)
CWRU	288	(0.7)	38	(1.9)
CoRIS	894	(2.0)	13	(0.7)
EuroSIDA	531	(1.2)	13	(0.7)
FENWAY	292	(0.7)	10	(0.5)
FHDH	6852	(15.6)	286	(14.3)
Frankfurt	104	(0.2)	5	(0.3)
GEMES	196	(0.5)	7	(0.4)
HAVACS	200	(0.5)	9	(0.5)
ICONA	864	(2.0)	32	(1.6)
JH	470	(1.1)	36	(1.8)
PISCIS	1230	(2.8)	32	(1.6)
PRIMO	251	(0.6)	1	(0.1)
SEROCO	14	(0.0)	0	(0.0)
SHCS	1892	(4.3)	42	(2.1)
UAB	337	(0.8)	23	(1.2)
UCSD	637	(1.5)	44	(2.2)
UCSF	583	(1.3)	47	(2.4)
UKREG	692	(1.6)	37	(1.9)
UNC	450	(1.0)	43	(2.2)
VACH	3743	(8.6)	184	(9.2)
VACS	2231	(5.1)	146	(7.3)
Vanderbilt	451	(1.0)	42	(2.1)
Washington	361	(0.8)	19	(1.0)

* The full names of the cohorts are provided in Appendix 1.

Appendix Table 3: Key dates and HIV-1 RNAs for the 29 individuals contributing deaths in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

Strategy	Individual	Baseline		Crosses Tight Threshold		Crosses Loose Threshold		Switch Date	Death Date
		Date	RNA	Date	RNA	Date	RNA		
400 only	1	7/18/2006	404	7/18/2006	404			10/11/2006	7/27/2007
	2	6/22/2004	210	8/8/2005	477	.	.	.	9/25/2006
	3	5/25/2004	258	5/26/2005	838	.	.	.	3/21/2008
	4	3/12/2005	271	12/31/2005	744			.	6/2/2006
1000 only	5	10/21/2002	871	10/21/2002	871	.	.	.	2/11/2003
	6	1/20/2009	641	1/20/2009	641		.	.	7/13/2010
	7	11/16/2007	584	11/16/2007	584	.	.	.	1/15/2012
	8	11/16/2004	479	11/16/2004	479	.	.	.	9/8/2005
	9	9/28/2004	614	9/28/2004	614	9/6/2006	19076	9/21/2006	11/22/2006
	10	7/7/2004	254	.	.	.	.	.	3/8/2005
	11	2/15/2005	220	.	.	.	.	.	3/21/2007
	12	11/13/2008	366	.	.	.	.	.	2/11/2009
	13	4/7/2005	283	.	.	.	.	.	7/17/2005
	14	6/4/2008	400	.	.	.	.	.	11/15/2008
	15	5/14/2007	400	.	.	.	.	.	3/15/2008
Both 400 and 1000	16	5/17/2007	100010	5/17/2007	100010	5/17/2007	100010	.	5/27/2007
	17	6/15/2005	165797	6/15/2005	165797	6/15/2005	165797	.	9/9/2005
	18	6/13/2006	54875	6/13/2006	54875	6/13/2006	54875	.	8/27/2006
	19	4/6/2006	1754	4/6/2006	1754	4/6/2006	1754	.	5/6/2006
	20	4/11/2005	694000	4/11/2005	694000	4/11/2005	694000	.	4/29/2005
	21	10/20/2009	427233	10/20/2009	427233	10/20/2009	427233	11/2/2009	6/7/2010
	22	12/14/2004	54172	12/14/2004	54172	12/14/2004	54172	1/11/2005	12/10/2005

23	11/3/2009	3674	11/3/2009	3674	11/3/2009	3674	12/15/2009	4/7/2010
24	3/9/2006	28000	3/9/2006	28000	3/9/2006	28000	3/9/2006	5/16/2009
25	3/21/2007	100000	3/21/2007	100000	3/21/2007	100000	4/17/2007	3/15/2012
26	1/13/2006	1019	1/13/2006	1019	1/13/2006	1019	1/13/2006	1/15/2007
27	2/17/2004	30705	2/17/2004	30705	2/17/2004	30705	3/3/2004	8/20/2005
28	12/6/2004	39896	12/6/2004	39896	12/6/2004	39896	2/14/2005	5/23/2008
29	7/6/2006	246	10/19/2006	1207	10/19/2006	1207	10/19/2006	2/5/2007

Appendix Table 4: Hazard ratios of clinical outcomes for tight and loose control switching strategies in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012: sensitivity analysis when no confirmation of virologic failure is required for eligibility.

Outcome	Strategy (HIV-1 RNA threshold in copies/mL)	No. of outcomes (overlap with tight)	Median (IQR) time to event in weeks	Hazard Ratio, 95% confidence interval					
				Unadjusted		Baseline-adjusted		Baseline and time-varying adjusted* †	
Death	Tight (400)	50	34 (8, 80)	1 (ref.)		1 (ref.)		1 (ref.)	
	Loose (1,000)	63 (47)	37 (9, 87)	1.01	0.86, 1.17	1.01	0.88, 1.17	0.86	0.55, 1.33
AIDS or death	Tight (400)	6860	9 (3, 46)	1 (ref.)		1 (ref.)		1 (ref.)	
	Loose (1,000)	83 (66)	19 (5, 57)	1.00	0.90, 1.12	0.99	0.88, 1.10	1.06	0.79, 1.43

Abbreviations: HIV, human immunodeficiency virus; IQR, interquartile range; Ref, reference

* Adjusted for the baseline covariates (sex, age, race, geographic origin, mode of acquisition, CD4 cell count, HIV RNA, calendar year, regimen class at initiation, and regimen class at baseline) and time-varying covariates (CD4 cell count, HIV RNA, AIDS-defining illness, and time since last laboratory measurement).

† Time-varying adjustment carried out by inverse probability weighting with weights truncated at the 99th percentile.

Figure 1: Modified CONSORT flow diagram for the mortality analysis in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.

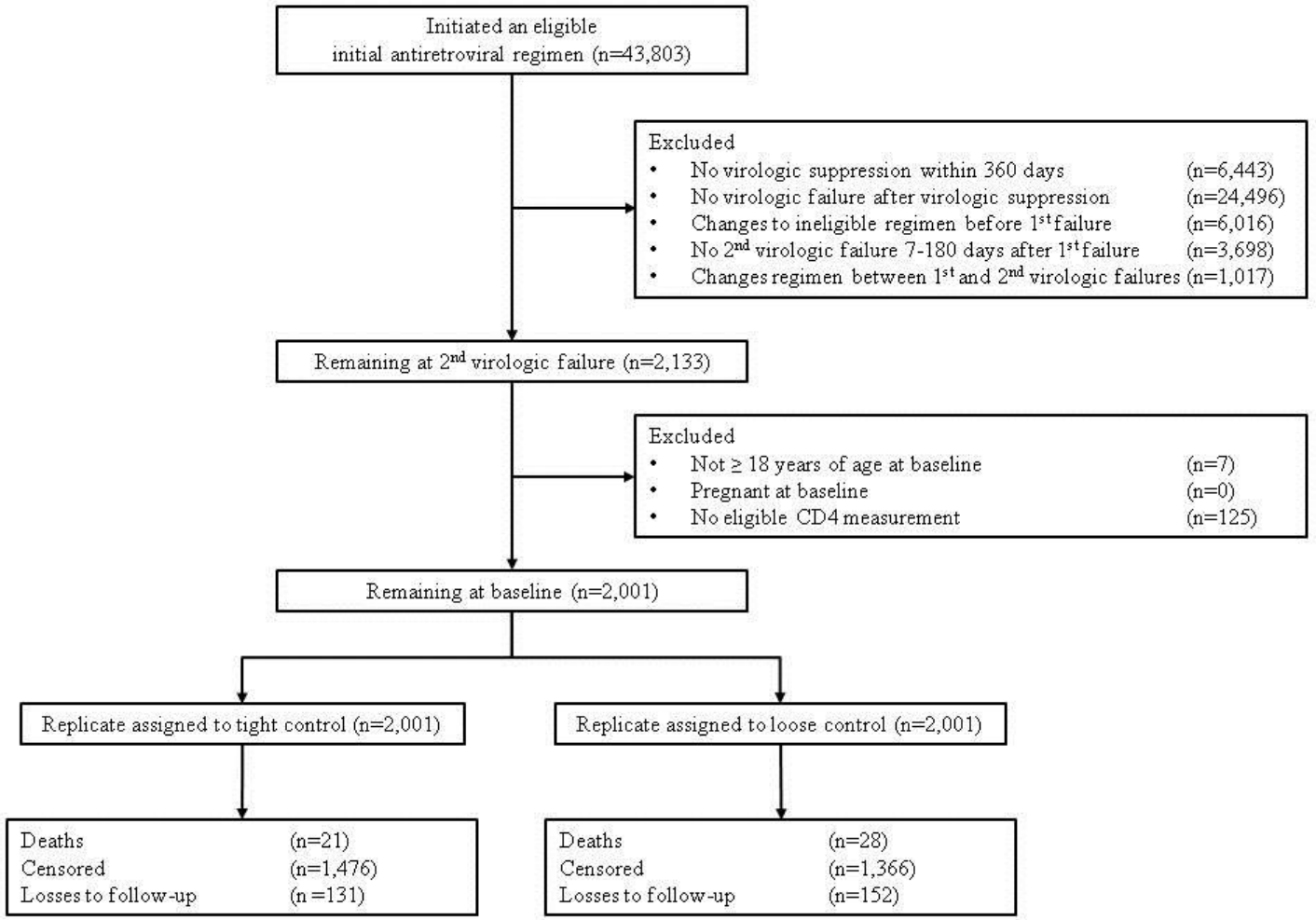
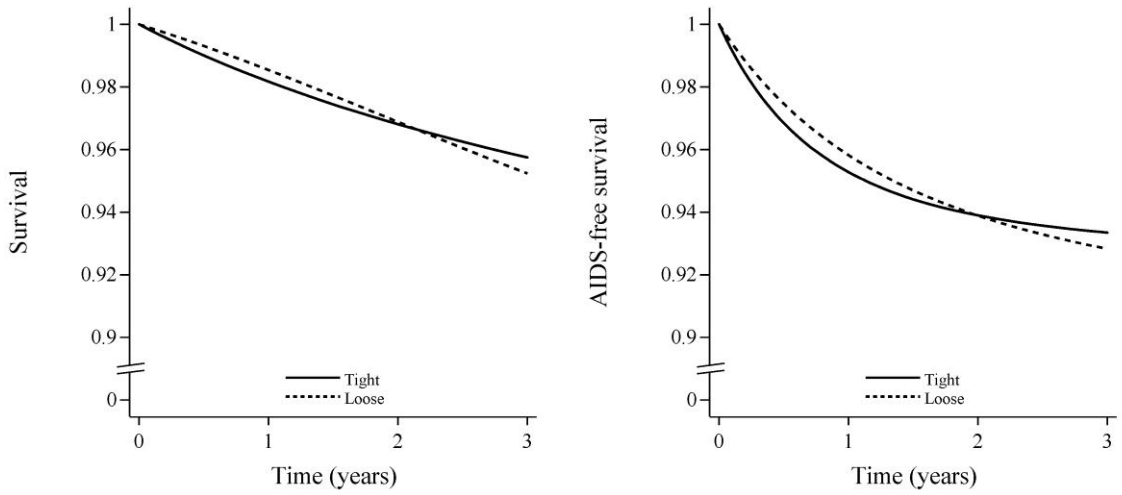
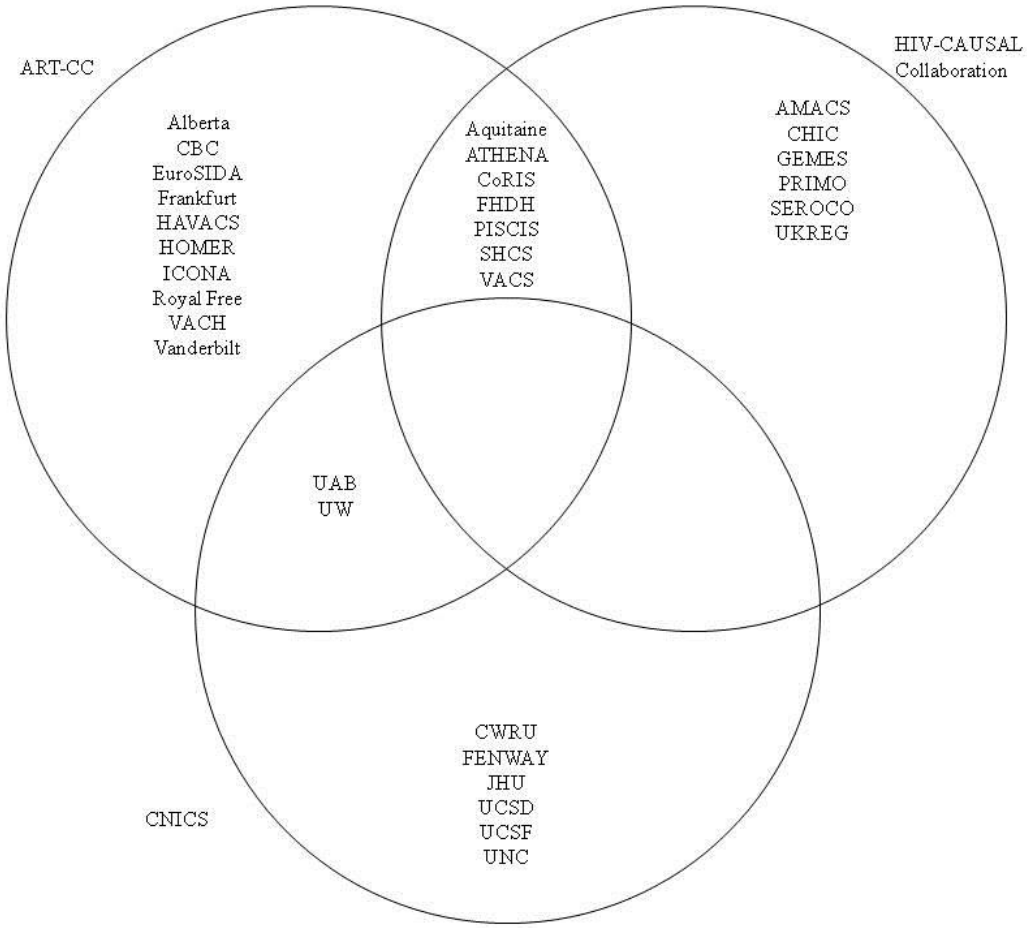


Figure 2: Survival (left) and AIDS-free survival (right) under tight and loose control switching strategies in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.



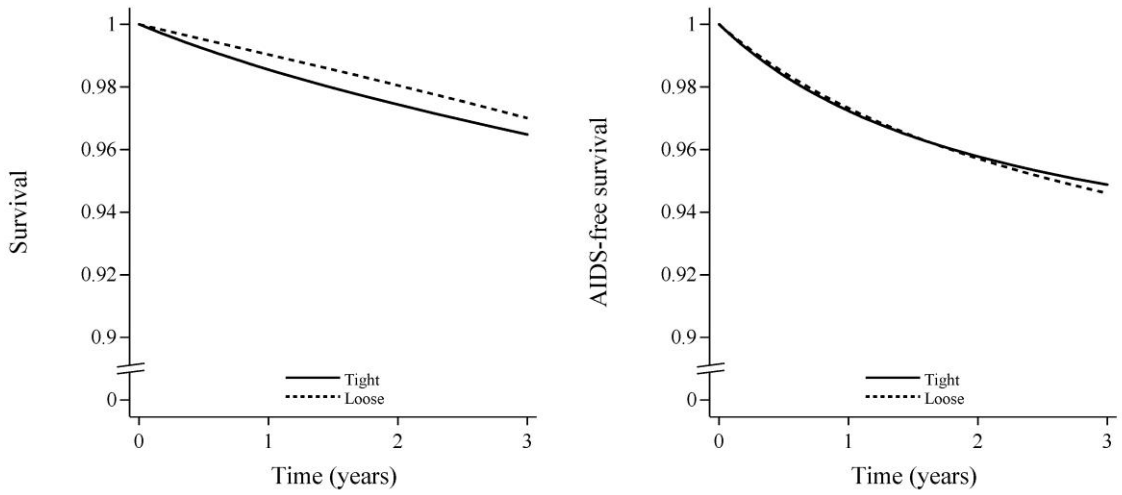
*The curves are standardized by the baseline covariates and inverse probability-weighted by the time-varying covariates listed under Table 4.

Appendix Figure 1: Venn Diagram of cohorts participating the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012.



* The full names of the cohorts are provided in Appendix 1.

Appendix Figure 2: Survival (left) and AIDS-free survival (right) under tight and loose control switching strategies in the ART-CC, the CNICS, and the HIV-CAUSAL Collaboration, 2002-2012: sensitivity analysis when no confirmation of virologic failure is required for eligibility.



*The curves are standardized by the baseline covariates and inverse probability-weighted by the time-varying covariates listed under Table 5.

REFERENCES

1. Panel on Antiretroviral Guidelines for Adults and Adolescents. *Guidelines for the use of antiretroviral agents in HIV-1 infected adults and adolescents*: Department of Health and Human Services. Available at http://www.aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf; 2014.

2. Gunthard HF, Aberg JA, Eron JJ, et al. Antiretroviral treatment of adult HIV infection: 2014 recommendations of the International Antiviral Society-USA Panel. *JAMA* 2014; **312**: 410-25.

3. European AIDS Clinical Society. *EACS Guidelines*: Available at http://www.eacsociety.org/Portals/0/Guidelines_Online_131014.pdf; 2013.

4. Writing G, Williams I, Churchill D, et al. British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012 (Updated November 2013. All changed text is cast in yellow highlight.). *HIV Med* 2014; **15 Suppl 1**: 1-85.

5. World Health Organization. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727_eng.pdf?ua=1. 2013.

6. Hatano H, Hunt P, Weidler J, et al. Rate of viral evolution and risk of losing future drug options in heavily pretreated, HIV-infected patients who continue to receive a stable, partially suppressive treatment regimen. *Clin Infect Dis* 2006; **43**: 1329-36.

7. Petersen ML, van der Laan MJ, Napravnik S, Eron JJ, Moore RD, Deeks SG. Long-term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification. *AIDS* 2008; **22**: 2097-106.

8. Johnson BA, Ribaudo H, Gulick RM, Eron JJ, Jr. Modeling clinical endpoints as a function of time of switch to second-line ART with incomplete data on switching times. *Biometrics* 2013; **69**: 732-40.

9. Li L, Eron JJ, Ribaudo H, Gulick RM, Johnson BA. Evaluating the Effect of Early Versus Late ARV Regimen Change if Failure on an Initial Regimen: Results From the AIDS Clinical Trials Group Study A5095. *Journal of the American Statistical Association* 2012; **107**: 542-54.

10. Hernan MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008; **19**: 766-79.

11. Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernán MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat* 2010; **6**: Article 18.

12. Garcia-Albeniz X, Logan RW, Schrag D, Hernan MA. Evaluation of the duplication of staging CT scans for localized colon cancer in a Medicare population. *Med Care* 2014; **52**: 963-8.

13. Danaei G, Rodriguez LA, Cantero OF, Logan R, Hernan MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical methods in medical research* 2013; **22**: 70-96.

14. Zhang Y, Thamer M, Kaufman J, Cotter D, Hernan MA. Comparative effectiveness of two anemia management strategies for complex elderly dialysis patients. *Med Care* 2014; **52 Suppl 3**: S132-9.

15. Robins J, Hernán M. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. *Longitudinal Data Analysis*. New York: Chapman and Hall/CRC Press; 2008. p. 553-99.

16. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period: application to the healthy worker survivor effect. *Mathematical Modeling* 1986; **7**: 1393-512.

17. van der Laan MJ, Petersen ML. Causal Effect Models for Realistic Individualized Treatment and Intention to Treat Rules. *Int J Biostat* 2007; **3**: Article3.

18. CDC. 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR* 1992; **41**: 1-19.

19. Hernan MA. The hazards of hazard ratios. *Epidemiology* 2010; **21**: 13-5.

20. Thompson WA, Jr. On the treatment of grouped observations in life studies. *Biometrics* 1977; **33**: 463-70.

21. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000; **11**: 561-70.

22. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic & Clinical Pharmacology & Toxicology* 2006; **98**: 237-42.

23. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main Content. *The International Journal of Biostatistics* 2010; **6**: Article 7.

24. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part II: Proofs and Additional Results. *The International Journal of Biostatistics* 2010; **6**: Article 8.

25. Little R.J., D'Agostino R., Cohen M.L., et al. The prevention and treatment of missing data in clinical trials. *The New England journal of medicine* 2012; **367**: 1355-60.

26. May MT, Ingle SM, Costagliola D, et al. Cohort profile: Antiretroviral Therapy Cohort Collaboration (ART-CC). *International journal of epidemiology* 2014; **43**: 691-702.

27. Egger M, May M, Chene G, et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet* 2002; **360**: 119-29.

28. The HIV-CAUSAL Collaboration, Ray M, Logan R, et al. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. *AIDS* 2010; **24**: 123-37.

29. Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. *Int J Epidemiol* 2008; **37**: 948-55.

30. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008; **168**: 656-64.

31. Garcia-Albeniz X, Chan JM, Paciorek A, et al. Immediate versus deferred initiation of androgen deprivation therapy in prostate cancer patients with PSA-only relapse. An observational follow-up study. *European journal of cancer* 2015.

32. Robins JM, Hernán MA, U S. Effects of multiple interventions. In: Ezzati M, Lopez AD, Rodgers A, CJL M, editors. *Comparative quantification of health risks: global and regional burden*

of disease attributable to selected major risk factors. World Health Organization, Geneva; 2004.

33. Taubman SL, Robins JM, Mittleman MA, Hernán MA. *Alternative approaches to estimating the effects of hypothetical interventions*. JSM Proceedings, Health Policy Statistics Section; 2008; Alexandria, VA: American Statistical Association; 2008.

34. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Statistics in Biosciences* 2011; **3**: 119-43.