Title

Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy

Permalink

https://escholarship.org/uc/item/3d5498fg

Journal

International Journal of Epidemiology, 45(6)

ISSN

0300-5771

Authors

Cain, Lauren E
Saag, Michael S
Petersen, Maya
et al.

Publication Date

2016-12-01

DOI

10.1093/ije/dyv295

Peer reviewed

Methodological insights

Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy

Writing committee: Lauren E Cain,[1]* Michael S Saag,[2] Maya Petersen,[3] Margaret T May,[4] Suzanne M Ingle,[4] Roger Logan,[1] James M Robins,[1,5] Sophie Abgrall,[6,7] Bryan E Shepherd,[8] Steven G Deeks,[9] M John Gill,[10] Giota Touloumi,[11] Georgia Vourli,[11] François Dabis,[12] Marie-Anne Vandenhende,[12] Peter Reiss,[13,14] Ard van Sighem,[13] Hasina Samji,[15] Robert S Hogg,[15,16] Jan Rybniker,[17] Caroline A Sabin,[18] Sophie Jose,[18] Julia del Amo,[19,20] Santiago Moreno,[21,22] Benigno Rodríguez,[23] Alessandro Cozzi-Lepri,[24] Stephen L Boswell,[25] Christoph Stephan,[26] Santiago Pérez-Hoyos,[27] Inma Jarrin,[19,20] Jodie L Guest,[28,29,30] Antonella D'Arminio Monforte,[31] Andrea Antinori,[32] Richard Moore,[33] Colin NJ Campbell,[20,34] Jordi Casabona,[20,34,35] Laurence Meyer,[36] Rémonie Seng,[36] Andrew N Phillips,[18] Heiner C Bucher,[37] Matthias Egger,[38,39] Michael J Mugavero,[40] Richard Haubrich,[41] Elvin H Geng,[42] Ashley Olson,[43] Joseph J Eron,[44] Sonia Napravnik,[44] Mari M Kitahata,[45] Stephen E Van Rompaey,[45] Ramón Teira,[46] Amy C Justice,[47,48] Janet P Tate,[47,48] Dominique Costagliola,[49] Jonathan AC Sterne,[4] and Miguel A Hernán[1,5,50] on behalf of the Antiretroviral Therapy Cohort Collaboration, the Centers for AIDS Research Network of Integrated Clinical Systems, and the HIV-CAUSAL Collaboration

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [2]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA, [3]Divisions of Biostatistics and Epidemiology, University of California, Berkeley, School of Public Health, Berkeley, CA, USA, [4]School of Social and Community Medicine, University of Bristol, Bristol, UK, [5]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [6]Sorbonne Universités, UPMC Univ Paris 06, INSERM, Institut Pierre Louis d'épidémiologie et de Santé Publique (IPLESP UMRS 1136), F75013, Paris, France, [7]Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Antoine Béclère, Service de Médecine Interne, Clamart, France, [8]Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA, [9]Positive Health Program, San Francisco General Hospital, San Francisco, CA, USA, [10]Division of Infectious Diseases, University of Calgary, Calgary, AB, Canada, [11]Department of Hygiene, Epidemiology and Medical Statistics, Athens University Medical School, Athens, Greece, [12]INSERM U897, Centre Inserm Epidémiologie et Biostatistique, Université de Bordeaux, and Bordeaux University Hospital,

Department of Internal Medicine, Bordeaux, France, [13]Stichting HIV Monitoring, Amsterdam, Netherlands, [14]Academic Medical Center, Department of Global Health and Division of Infectious Diseases, University of Amsterdam, and Amsterdam Institute for Global Health and Development, Amsterdam, Netherlands, [15]Epidemiology and Population Health Program, BC Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada, [16]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada, [17]1st Department of Internal Medicine, University of Cologne, D-50937 Cologne, Germany, [18]University College London, London, UK, [19]National Centre of Epidemiology, Instituto de Salud Carlos III, Madrid, Spain, [20]Consorcio de Investigación Biomédica de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain, [21]Ramón y Cajal Hospital, IRYCIS, Madrid, Spain, [22]University of Alcalá de Henares, Madrid, Spain, [23]Division of Infectious Disease, Case Western Reserve University, Cleveland, OH, USA, [24]Department of Infection and Population Health; Division of Population Health, University College London, London, UK, [25]Fenway Health, Boston, MA, USA, [26]HIV Center, Department of Infectious Diseases, University Hospital, Frankfurt, Germany, [27]Vall d'Hebron Research Institute, Barcelona, Spain, [28]Rollins School of Public Health at Emory University, Atlanta, GA, USA, [29]Emory University School of Medicine, Atlanta, GA, USA, [30]Atlanta Veterans Affairs Medical Center, Decatur, GA, USA, [31]Clinic of Infectious Diseases and Tropical Medicine, Department of Health Sciences, San Paolo Hospital, University of Milan, Milan, Italy, [32]Istituto Nazionale per le Malattie Infettive Lazzaro Spallanzani IRCCS, Rome, Italy, [33]School of Medicine, Johns Hopkins University, Baltimore, MD, USA, [34]Center for Epidemiological Studies on HIV/AIDS and STI of Catalonia (CEEISCAT), Agència Salut Pública de Catalunya (ASPC), Generalitat de Catalunya, Badalona, 08916 Catalonia, Spain, [35]Department of Paediatrics, Obstetrics, Gynaecology and Preventive Medicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Catalonia, Spain, [36]Université Paris Sud, INSERM CESP U1018, and AP-HP, Hôpital de Bicêtre, Service de Santé Publique, le Kremlin Bicêtre, France, [37]Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel, Switzerland, [38]Centre for Infectious Disease Epidemiology and Research, School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, [39]University of Bern, Institute for Social and Preventive Medicine, Bern, Switzerland, [40]University of Alabama at Birmingham, Birmingham, AL, USA, [41]University of California San Diego, CA, USA (Currently Gilead Sciences, Foster City, CA, USA), [42]Division of HIV/AIDS, Department of Medicine, University of California, San Francisco, CA, USA, [43]Medical Research Council Clinical Trials Unit, University College London, London, UK, [44]Division of Infectious Diseases, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [45]Department of Medicine, University of Washington, Seattle, WA, USA, [46]Unit of Infectious Diseases, Hospital Sierrallana, Torrelavega, Spain, [47]Yale School of Medicine, New Haven, CT, USA, [48]VA Connecticut Healthcare System, West Haven, CT, USA, [49]Sorbonne Universités, UPMC Univ Paris 06, INSERM, Institut Pierre Louis d'épidémiologie et de Santé Publique (IPLESP UMRS 1136), F75013, Paris, France and [50]Harvard-MIT Division of Health Sciences and Technology, Boston, MA, USA

*Corresponding author. Department of Epidemiology, Harvard T. H. Chan School of Public Health, 677 Huntington Avenue, Boston MA 02115, USA. E-mail: lcain@hsph.harvard.edu

Abstract

Background: When a clinical treatment fails or shows suboptimal results, the question of when to switch to another treatment arises. Treatment switching strategies are often dynamic because the time of switching depends on the evolution of an individual's time-varying covariates. Dynamic strategies can be directly compared in randomized trials. For example, HIV-infected individuals receiving antiretroviral therapy could be 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: We review an approach to emulate a randomized trial of dynamic switching strategies using observational data from the Antiretroviral Therapy Cohort Collaboration, the Centers for AIDS Research Network of Integrated Clinical Systems and the

HIV-CAUSAL Collaboration. We estimated the comparative effect of tight-control vs. loose-control strategies on death and AIDS or death via inverse-probability weighting.

Results: Of 43 803 individuals who initiated an eligible antiretroviral therapy regimen in 2002 or later, 2001 met the baseline inclusion criteria for the mortality analysis and 1641 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 adjusted hazard ratios (95% confidence interval) for loose control were 1.10 (0.73, 1.66) for death, and 1.04 (0.86, 1.27) for AIDS or death.

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

Key words: HIV, antiretroviral therapy, inverse-probability weighting, observational studies, mortality, dynamic strategies

---

Key Messages

- A randomized trial comparing dynamic treatment-switching 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 and analysis plan.
- This approach detected little or no differences between switching at HIV-1 RNA thresholds of 400 and 1000 copies/ml in preventing death and AIDS-defining illness.

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 dynamic strategies for switching 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 virological failure (i.e. two consecutive HIV-1 RNA measurements ≥ 200 copies/ml). The European AIDS Clinical Society[3] and British HIV Association[4] guidelines recommend switching immediately if HIV-1 RNA > 500 copies/ml or > 400 copies/ml, respectively, but suggest repeating a viral load measurement if HIV-1 RNA is detected below the threshold for switching. The World Health Organization[5] guidelines recommend waiting to switch until confirmation of HIV-1 RNA > 1000 copies/ml, the lowest level that can be used when measuring viral load from dried blood spots. Generally, the harder to implement tight-control strategies

are recommended to avoid the use of failing antiretrovirals in the presence of ongoing viral replication which may lead to drug-resistant mutations, use of more expensive drugs and limitations on 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. Discussions between colleagues with clinical and statistical backgrounds are generally needed to specify a target trial whose results would be most useful to resolve uncertainties in clinical practice. We describe the components of the target trial 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 virological failure (defined as the second of two measurements of HIV-1 RNA > 200 copies/ml 7–180 days apart). At confirmed virological failure (baseline), individuals are required to be 18 years of age or older and to have had a CD4 cell count measurement in the previous 90 days. Eligible antiretroviral regimens before first virological failure are listed in Table 1.

Treatment strategies

Eligible individuals are randomized to either tight- or loose-control strategies at confirmed virological 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 initial switch, individuals may subsequently switch to another regimen or discontinue treatment if clinically indicated or recommended by their treating

Table 1. Eligible initial regimens

Regimen classification	Eligible initial regimens[a]
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[b]
bPI + ≥ 2 NRTI	All regimens except those containing an excluded drug[b]
NNRTI + ≥ 2 NRTI	All regimens except those containing an excluded drug[b]
< 6 drugs including FI/INSTI (+ entry inhibitors)	All drug regimens with ≥ 3 drugs except those containing an excluded drug[b]

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.

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

[b]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 to new regimens that are considered switches[a]

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[b]	No	Yes	Yes
bPI + ≥ 2 NRTI	Yes	Yes if PI changes[b]	Yes	Yes
NNRTI + ≥ 2 NRTI	Yes	Yes	Yes if NNRTI to etravirine (ETV)[b]	Yes
bPI + PI/NNRTI (+ other)	Yes	Yes	Yes if NNRTI to etravirine (ETV)[b]	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[b]

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.

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).

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.

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

[b]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).

physicians. 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. As in all randomized trials, we expect that some individuals will not adhere to the target trial protocol.

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 estimate 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 $\Pr(D_{t+1} = 1 \mid D_t = 0, X) = \beta_{0t} + \beta_1 X$, 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 vs 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 their assigned switching strategy. 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 a per-protocol mortality hazard ratio, we fit a weighted model logit $\Pr(D_{t+1} = 1 \mid D_t = 0, C_t = 0, X, V) = \theta_{0t} + \theta_1 x + \theta_2' V$, 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), geographical 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, 401–1000, > 1000 copies/ml), calendar year (2002–04, 2005–07, ≥ 2008), regimen class at initiation [non-nucleoside (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 time-varying 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 ($\overline{L}_t$), and his/her switching history ($\overline{A}_{t-1}$), where $A_t = 2$ indicates that the individual has switched to an eligible regimen, $A_t = 1$ indicates that the individual has changed to an ineligible regimen, 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: $W_t = \prod_{k=0}^{t} \frac{1}{f(A_k \mid \overline{A}_{k-1}, D_k = 0, \overline{L}_k)}$, where $f(A_k \mid \overline{A}_{k-1}, D_k = 0, \overline{L}_k)$ is the conditional probability mass function $f_{A_k \mid \overline{A}_{k-1}, D_k = 0, \overline{L}_k}(a_k \mid \overline{a}_{k-1}, d_k = 0, \overline{l}_k)$ with $(a_k \mid \overline{a}_{k-1}, d_k = 0, \overline{l}_k)$ evaluated at the random argument $(A_k \mid \overline{A}_{k-1}, D_k = 0, \overline{L}_k)$ 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 1000 cells/mm^3), HIV-1 RNA (≤ 400, 401–1000, >1000 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: (i) we measured and successfully adjusted for all baseline and time-varying confounders; (ii) there is positivity (i.e. no deterministic treatment assigned given the confounders); and (iii) the weight models are not misspecified, the above logistic model consistently estimates the parameter of a dynamic marginal logistic structural model:[17,22–24]
logit $\Pr\left(D_{t+1}^x = 1 \mid D_t^x = 0, V\right) = \theta_{0t}^* + \theta_1^* x + \theta_2^{*'} V$, 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 model's 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 analytical 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 the 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 (available as Supplementary data at *IJE* online). All overlaps between and within collaborations were removed. Each cohort collected data prospectively, including CD4 cell count, HIV-1 RNA (limit of detection $\leq$ 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 1 January 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 of 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 dataset constructed as follows. Because all individuals had data consistent with both strategies at confirmed virological 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
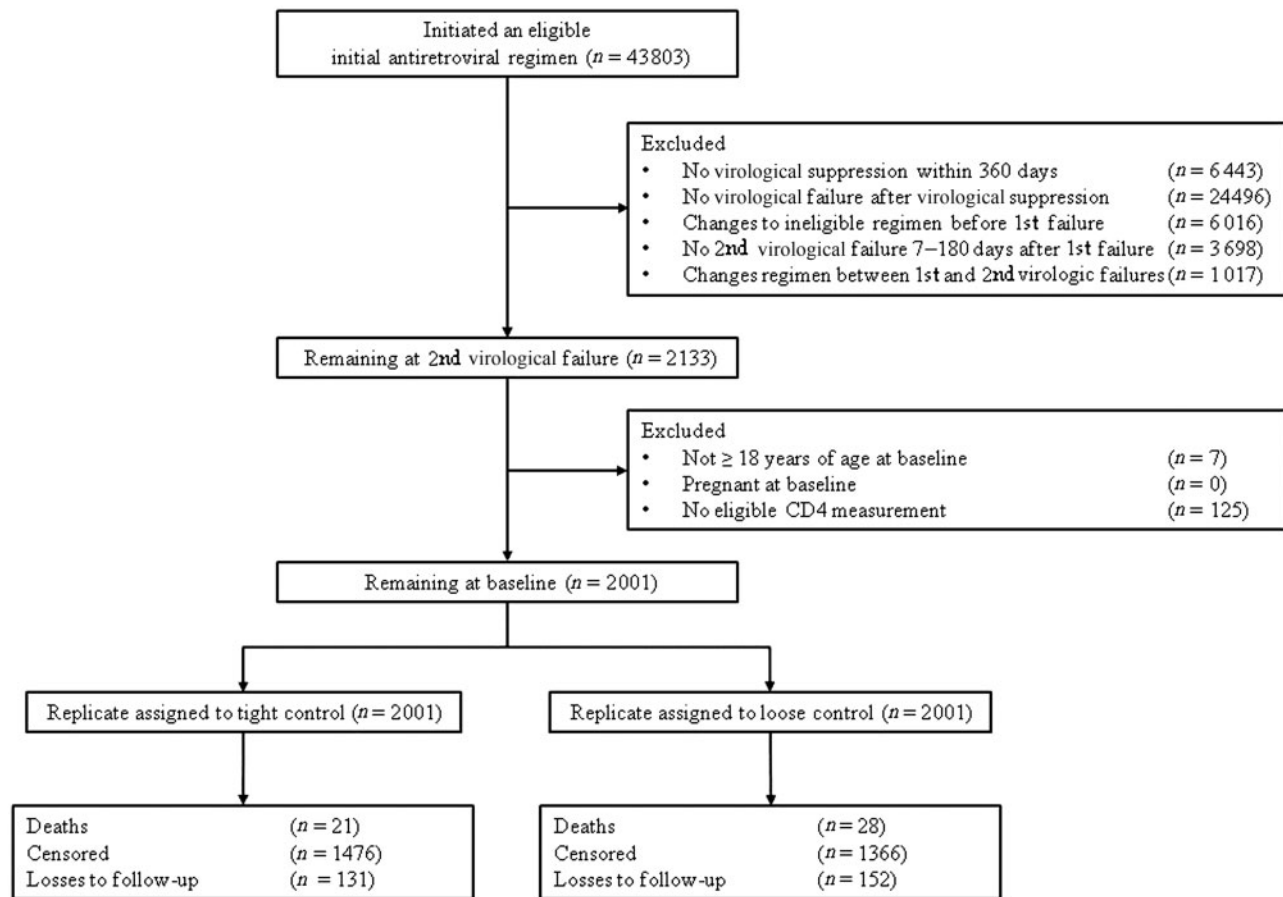
Figure 1. Modified CONSORT flow diagram for the mortality analysis in the ART-CC, the CNICS and the HIV-CAUSAL Collaboration, 2002–12.

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 $0 \le j < m$ if the individual initiates, and $\left(1 - \frac{1}{m+1-j}\right)$ when $0 \le j < m$ 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 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 (data not shown).

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

Results

Of 43 803 potentially eligible individuals, 2001 met the baseline inclusion criteria for the mortality analysis and 1641 for the AIDS or death analysis. The most common reason for being excluded was never experiencing virological failure after achieving virological 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 4002

Table 3. Characteristics of 2001 HIV-infected individuals in the mortality analysis at baseline in the ART-CC, the CNICS and the HIV-CAUSAL Collaboration, 2002–12

Characteristic		No. of individuals (%)			
		Western Europe		North America	
		1503		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)
Geographical Origin	North America	0	(0.0)	498	(100.0)
	Western Europe	587	(39.1)	0	(0.0)
	Sub-Saharan Africa	516	(34.3)	0	(0.0)
	Other/unknown	400	(26.6)	0	(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[a]	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)
	>1000	916	(60.9)	256	(51.4)
Calendar year	2002–04	256	(17.0)	157	(31.5)
	2005–07	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)

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.

[a]Other/unknown acquisition group included all VACS-VC participants.

replicates in the expanded dataset, 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, the respective numbers were 23%, 14% and 63%. Among the uncensored, the median [interquartile range (IQR)] follow-up time was 89

(38, 168) weeks for the tight-control group (1673 person-years) and 82 (40, 166) weeks for the loose-control group (2009 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, available as Supplementary data at *IJE* online). Among those who died, the median (IQR) time to death was 31 (11, 52) weeks for the tight-control
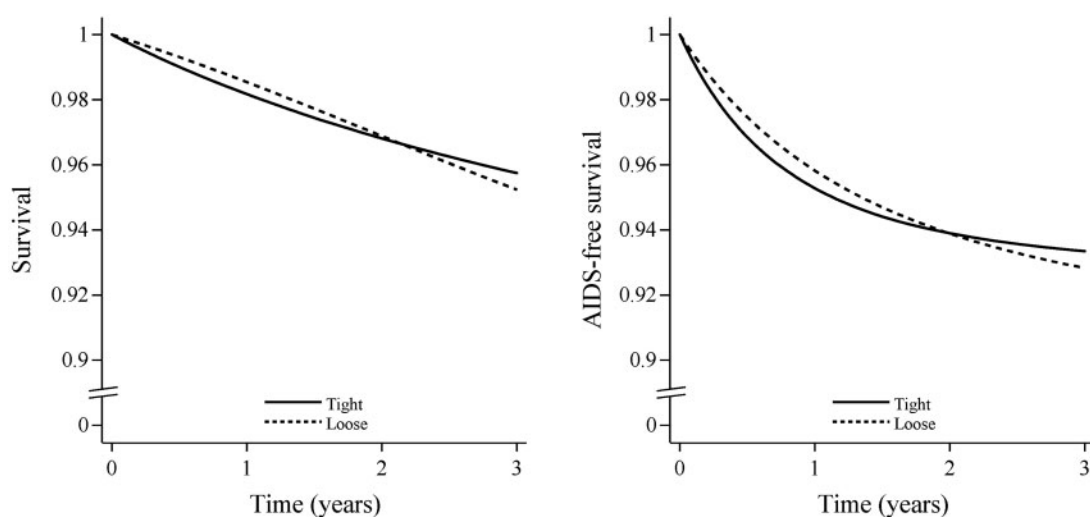
Table 4. Hazard ratios of clinical outcomes under tight- and loose-control switching strategies in the ART-CC, the CNICS and the HIV-CAUSAL Collaboration, 2002–12

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[a,b]	
Death	Tight (400)	21	31 (11, 52)	1 (ref.)		1 (ref.)		1 (ref.)	
	Loose (1000)	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 (1000)	41 (31)	15 (7, 60)	1.08	0.90, 1.28	1.05	0.90, 1.23	1.04	0.86, 1.27

HIV, human immunodeficiency virus; IQR, interquartile range; ref., reference.

[a]Adjusted for the baseline covariates (sex, age, race, geographical 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).

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

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, 2002–12. The curves are standardized by the baseline covariates and inverse-probability weighted by the time-varying covariates listed under Table 4.

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 4). The estimated inverse-probability weights had mean 3.1 (interquartile range 1.2, 3.2, 99th percentile 15.5) for the mortality analysis, and 3.1 (interquartile range 1.1 − 3.4, 99th percentile 17.2) for the AIDS or death analysis. The main predictors of switching were HIV-1 RNA and time since last laboratory measurement (see Appendix Table 1, available as Supplementary data at *IJE* online).

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 virological failure. In this case, baseline becomes the time of first virological failure (defined as one measurement of HIV-1 RNA > 200 copies/ml) following virological suppression. Estimated hazard ratios using this definition of baseline were similar (see Appendix 3 for details, available as Supplementary data at *IJE* online).

Discussion

We have described how to use observational data to emulate a target trial that compares 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 virological failure.

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 virological failure. Had we been able to observe individuals for longer periods of time, more of them would likely have experienced virological failure and could have been included in our analyses.

Most individuals in our analyses contributed to both arms of the target trial. Because one cannot generally observe the exact moment at which HIV-1 RNA thresholds were crossed, 59% of individuals crossed both thresholds at baseline (those with baseline HIV-RNA ≤ 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 previous 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. The dates of pregnancies were not available for the majority of individuals in this analysis, and 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. Data on adherence and the results of resistance testing were not available for the majority of individuals in this analysis, but 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, which may result in more precise estimates at the expense of additional modelling assumptions.[16,32–34]

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

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, 2015. http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf (8 June 2015, date last accessed).

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.* 2015. http://www.eacsociety.org/files/guidelines_8_0-english_web.pdf (22 October 2015, date last accessed).

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. 2013. http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727_eng.pdf?ua=1 (13 July 2015, date last accessed).

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. *J Am Stat Assoc* 2012;**107**:542–54.

10. Hernán 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, Hernán MA. Evaluation of the duplication of staging CT scans for localized colon cancer in a Medicare population. *Med Care* 2014;**52**:963–68.

13. Danaei G, Rodriguez LA, Cantero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods med Res* 2013;**22**:70–96.

14. Zhang Y, Thamer M, Kaufman J, Cotter D, Hernán MA. Comparative effectiveness of two anemia management strategies

for complex elderly dialysis patients. *Med Care* 2014;**52**(Suppl 3):S132–39.

15. Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). *Longitudinal Data Analysis*. New York, NY: Chapman and Hall/CRC Press, 2008.

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 Modelling* 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. Centers for Disease Control and Prevention. 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morb Mortal Wkly Rep* 1992;**41**:1–19.

19. Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010;**21**:13–15.

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 Clin Pharmacol Toxicol* 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. *Int J Biostat* 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. *Int J Biostat* 2010;**6**:Article 8.

25. Little RJ, D'Agostino R, Cohen ML *et al*. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;**367**:1355–60.

26. May MT, Ingle SM, Costagliola D *et al*. Cohort profile: Antiretroviral Therapy Cohort Collaboration (ART-CC). *Int J Epidemiol* 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. *Eur J Cancer* 2015;**51**:817–24.

32. Robins JM, Hernán MA, Siebert U. Effects of multiple interventions. In: Ezzati M, Lopez AD, Rodgers A, Murry CJL (eds). *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva: World Health Organization, 2004.

33. Taubman SL, Robins JM, Mittleman MA, Hernán MA. *Alternative Approaches to Estimating the Effects of Hypothetical Interventions*. 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. *Stat Biosci* 2011;**3**:119–43.

Commentary: Should the analysis of observational data always be preceded by specifying a target experimental trial?

Vanessa Didelez

School of Mathematics, University of Bristol, Bristol, UK. E-mail: Vanessa.Didelez@bristol.ac.uk

The recent paper by Cain *et al*.[1] deals with the important practical question of when to switch antiretroviral therapy for HIV-infected individuals after virological failure.

However, the real significance of the paper is in advocating, and demonstrating the feasibility of, what we may want to adopt as a general principle: analysing