Original Contribution

Effect of Highly Active Antiretroviral Therapy on Incident AIDS Using Calendar Period as an Instrumental Variable

Lauren E. Cain, Stephen R. Cole, Sander Greenland, Todd T. Brown, Joan S. Chmiel, Lawrence Kingsley, and Roger Detels

Human immunodeficiency virus (HIV) researchers often use calendar periods as an imperfect proxy for highly active antiretroviral therapy (HAART) when estimating the effect of HAART on HIV disease progression. The authors report on 614 HIV-positive homosexual men followed from 1984 to 2007 in 4 US cities. During 5,321 person-years, 268 of 614 men incurred acquired immunodeficiency syndrome, 49 died, and 90 were lost to follow-up. Comparing the pre-HAART calendar period (<1996) with the HAART calendar period (≥1996) resulted in a naive rate ratio of 3.62 (95% confidence limits: 2.67, 4.92). However, this estimate is likely biased because of misclassification of HAART use by calendar period. Simple calendar period approaches may circumvent confounding by indication at the cost of inducing exposure misclassification. To correct this misclassification, the authors propose an instrumental-variable estimator analogous to ones previously used for noncompliance corrections in randomized clinical trials. When the pre-HAART calendar period was compared with the HAART calendar period, the instrumental-variable rate ratio was 5.02 (95% confidence limits: 3.45, 7.31), 39% higher than the naive result. Weighting by the inverse probability of calendar period given age at seroconversion, race/ethnicity, and time since seroconversion did not appreciably alter the results. These methods may help resolve discrepancies between observational and randomized evidence.

acquired immunodeficiency syndrome; bias (epidemiology); causality; confounding factors (epidemiology); HIV

Abbreviations: AIDS, acquired immunodeficiency syndrome; CL, confidence limits; HAART, highly active antiretroviral therapy; HIV, human immunodeficiency virus; ITT, intent to treat.

Observational studies have shown a protective effect of highly active antiretroviral therapy (HAART) on time to acquired immunodeficiency syndrome (AIDS). The studies reported by Detels et al. (1) and Porter et al. (2) combine to yield an inverse-variance weighted summary hazard ratio of 2.89 (95% confidence limits (CL): 2.19, 3.79) comparing a pre-HAART calendar period with the HAART calendar period. These studies used calendar period as a proxy for actual HAART use to circumvent confounding by indication (3). To the extent that calendar period is a misclassified version of actual HAART exposure, estimates using calendar period may be subject to information bias (M. A. Hernán, Harvard School of Public Health, unpublished manuscript, 2009).

Randomized clinical trials have also shown a protective effect of HAART on time to AIDS or death. Trials reported by Hammer et al. (4) and Cameron et al. (5) combine to yield an inverse-variance weighted summary hazard ratio of 1.91 (95% CL: 1.57, 2.33) comparing a non-HAART regimen consisting of 2 nucleoside reverse transcriptase inhibitors with a HAART regimen consisting of the same 2 nucleoside reverse transcriptase inhibitors plus a protease inhibitor. Randomized clinical trials are generally analyzed by using an intent-to-treat (ITT) approach, which compares endpoints as randomly assigned, irrespective of postrandomization compliance (6, p. 16). In the presence of noncompliance, this ITT approach typically produces null-biased estimates of the difference between 2 treatment arms. Methods exist to correct ITT estimates for noncompliance (7–14). For example, correcting the result of Hammer et al. for noncompliance by using inverse probability-of-censoring

Correspondence to Dr. Lauren E. Cain, Department of Epidemiology, 677 Huntington Avenue, Boston, MA 02115 (e-mail: lcain@hsph.harvard.edu).

weights increased the hazard ratio from 1.98 (95% CL: 1.30, 3.01) to 2.21 (95% CL: 1.34, 3.66) (15).

Using observational data from the Multicenter AIDS Cohort Study (16), we adapted existing noncompliance correction methods (8, 17) to correct for information bias due to use of an imperfect instrument. We extended the method (8) to estimate the rate (i.e., number of events divided by person-time at risk) instead of the risk (i.e., number of events divided by number of people at risk) and to account for measured covariates using inverse probability weights (18).
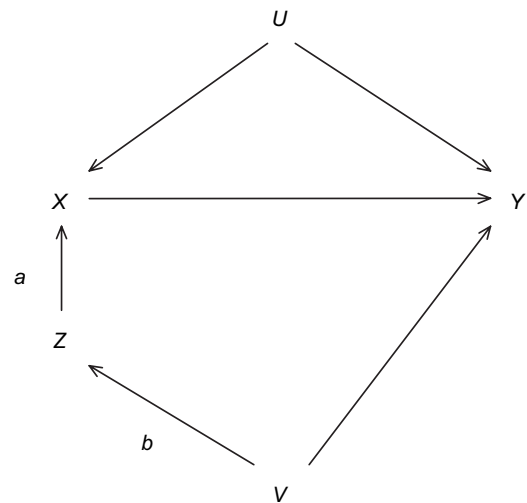
MATERIALS AND METHODS

Study population

The Multicenter AIDS Cohort Study is an ongoing prospective cohort study of the natural and treated history of human immunodeficiency virus (HIV) infection in 6,972 homosexual men enrolled starting in 1984 (16). Participants were seen semiannually in 4 US urban centers: Baltimore, Maryland/Washington, DC; Chicago, Illinois; Pittsburgh, Pennsylvania; and Los Angeles, California. During their study visit, they underwent a physical examination; provided biologic specimens, including a blood sample for CD4+ T-lymphocyte count, levels of plasma HIV RNA, and serologic HIV antibody tests on seronegative men; and completed an extensive interviewer- or computer-administered questionnaire. The questionnaire gathered information about medication use and medical history. Institutional review boards approved all protocols, and written informed consent was obtained for all participants. Our present analysis concerns the 614 men in the study who seroconverted between enrollment and April 2007.

Exposure assessment

Antiretroviral therapy use is measured by self-report using photo-medication cards. At each study visit, participants were asked about therapy use since their prior visit. We classify therapy use at each study visit as either HAART or non-HAART. The definition of HAART was guided by the DHHS/Kaiser Panel (19) guidelines, and HAART was defined as in the study by Cole et al. (20). Typical HAART regimens consisted of 2 or more nucleoside or nucleotide reverse transcriptase inhibitors in combination with at least 1 protease inhibitor or 1 nonnucleotide reverse transcriptase inhibitor. Combinations of zidovudine and stavudine with either a protease inhibitor or a nonnucleoside reverse transcriptase inhibitor were not considered HAART.

Each participant's person-time was partitioned into 2 calendar periods as defined by Detels et al. (1): pre-HAART (before 1996) and HAART (1996 and beyond). We also explored the impact of using 1998 as the cutoff for the HAART calendar period instead of 1996. An indicator variable for the HAART calendar period serves as a proxy for HAART use. This indicator appears to fulfill the 3 assumptions of an instrumental variable (17) and has been used in such a fashion previously (1, 2, 21–31). First, calendar periods are associated with therapy use because

Figure 1. Diagram showing the association of calendar period Z, therapy use X, AIDS Y, measured covariates V, and unmeasured covariates U. Refer to the Materials and Methods section of the text for more information. AIDS, acquired immunodeficiency syndrome.

therapies were introduced sequentially over time. In Figure 1, this statement corresponds to the existence of arrow a from calendar period Z to therapy use X. Second, calendar period Z cannot be affected by indications for treatment with therapy. In Figure 1, this statement is represented by the lack of an arrow from unmeasured confounders U to calendar period Z. Third, on the basis of the results of Detels et al., it appears that calendar period Z is independent of AIDS Y given indications for and actual use of therapy. In Figure 1, this statement corresponds to the absence of an arrow from calendar period Z to AIDS Y and the absence of variables V affecting both calendar period Z and AIDS Y. Detels et al. provide compelling evidence (in their Table 2) that HIV-related non-HAART therapies and health care utilization, 2 possible explanations for the apparent effect of HAART, do not differ notably across recent calendar periods.

Nonetheless, beyond indications for and actual use of therapy, because of the time course of the epidemic, some variables V may be associated with calendar period Z as well as being independent predictors of AIDS Y (e.g., age at seroconversion, race/ethnicity, and infection duration). Therefore, to relax the third assumption, we use inverse probability weighting to create a weighted pseudo-population in which there is no association of the variables V with calendar period Z, and hence no arrow b from variables V to calendar period Z. The observations in the weighted population are weighted by the inverse of the probability of calendar period given the measured variables V.

Endpoint ascertainment

The endpoint of interest was time from HIV seroconversion to clinical AIDS. The date of HIV seroconversion was

taken to be the midpoint between the date of a participant's last HIV-negative and first HIV-positive visits. The median interval between these 2 visits was 0.52 (interquartile range: 0.49, 0.71) years. The 1993 Centers for Disease Control and Prevention clinical conditions criteria were used to define clinical AIDS (32). Therefore, men with CD4 cell counts of less than 200 cells/mm^3 but no clinical conditions were not considered to have clinical AIDS. Reported cases of clinical AIDS were confirmed by using physician or hospital record abstractions (16). Men were censored at death, censored at loss to follow-up (defined as neither dead nor seen within the last 12 months of study), or administratively censored if still alive at the end of the study period in April 2007.

Statistical methods

The following notation will be used in this paper: subscript i indexes the 1 to $N = 614$ participants, and subscript j indexes the 1 to J_i visits for subject i. The maximum number of visits was 46. Subscript x indexes therapy use, where 1 is HAART use and 0 is non-HAART use. Subscript z indexes calendar period, where 1 is the HAART calendar period and 0 is the pre-HAART calendar period. $D_{ijxz} = 1$ indicates that participant i incurred the endpoint between visits j and $j + 1$ while using therapy x during calendar period z, 0 otherwise. T_{ijxz} is the number of person-years that participant i contributed between visits $j - 1$ and j while using therapy x during calendar period z. $\mathbf{V_{ij}}$ is a vector of time-fixed and time-varying covariates including age at seroconversion, race/ethnicity, and time since seroconversion.

Then, $D_{xz} = \sum_{i=1}^{614} \sum_{j=1}^{J_i} D_{ijxz}$ is the total number of events occurring while using therapy x during calendar period z summed over participants and visits. $T_{xz} = \sum_{i=1}^{614} \sum_{j=1}^{J_i} T_{ijxz}$ is the total number of person-years contributed while using therapy x during calendar period z summed over participants and visits. Let α_{xz} be the conditional probability of using therapy x given calendar period z, as estimated by the proportion of person-time spent using therapy x during calendar period z: $\alpha_{xz} = P(X = x \mid Z = z) = \frac{T_{xz}}{T_{+z}}$, where $T_{+z} = \sum_{x=0}^{1} T_{xz}$.

A standard ITT analysis compares rates between calendar periods, irrespective of actual therapy use. The ITT estimator of the average causal effect can be written as

$$\beta_{\text{ITT}} = \frac{\alpha_{10} \times \left(\frac{D_{10}}{T_{10}}\right) + \alpha_{00} \times \left(\frac{D_{00}}{T_{00}}\right)}{\alpha_{11} \times \left(\frac{D_{11}}{T_{11}}\right) + \alpha_{01} \times \left(\frac{D_{01}}{T_{01}}\right)} = \frac{\frac{D_{+0}}{T_{+0}}}{\frac{D_{+1}}{T_{+1}}}.$$

Our proposed instrumental-variable analysis compares rates between calendar periods among "compliers," those who would have used HAART in the HAART calendar period and would have used non-HAART in the pre-HAART calendar period. The instrumental-variable estimator, assuming

that calendar period is an appropriate instrumental variable and that calendar periods are exchangeable, can be written as

$$
\begin{aligned}
\beta_{\text{IV}} &= \frac{\alpha_{00} \times \left(\frac{D_{00}}{T_{00}}\right) - \alpha_{01} \times \left(\frac{D_{01}}{T_{01}}\right)}{\alpha_{11} \times \left(\frac{D_{11}}{T_{11}}\right) - \alpha_{10} \times \left(\frac{D_{10}}{T_{10}}\right)} \\
&= \frac{\beta_{\text{ITT}}}{\dfrac{\left[\alpha_{00} \times \left(\frac{D_{00}}{T_{00}}\right) - \alpha_{01} \times \left(\frac{D_{01}}{T_{01}}\right)\right] \times \left[\alpha_{11} \times \left(\frac{D_{11}}{T_{11}}\right) + \alpha_{01} \times \left(\frac{D_{01}}{T_{01}}\right)\right]}{\left[\alpha_{11} \times \left(\frac{D_{11}}{T_{11}}\right) - \alpha_{10} \times \left(\frac{D_{10}}{T_{10}}\right)\right] \times \left[\alpha_{10} \times \left(\frac{D_{10}}{T_{10}}\right) + \alpha_{01} \times \left(\frac{D_{01}}{T_{01}}\right)\right]}}.
\end{aligned}
$$

The ratio on the right-hand side shows β_{IV} in the common instrumental-variable format (17), with the ITT estimator divided by an estimator of the association between the instrumental variable and the exposure of interest. When there is perfect compliance (i.e., no one contributes person-time to the HAART calendar period while not using HAART, and no one contributes person-time to the pre-HAART calendar period while using HAART), then $\alpha_{10} = \alpha_{01} = 0$ and the equation above reduces to $\beta_{\text{IV}} = \beta_{\text{ITT}}$. In Appendix 1 we present a simulation comparison of the above ITT and instrumental-variable estimators.

To control for $\mathbf{V_{ij}}$, we build inverse probability of calendar period weights. The weights W_{ij+z} are defined as $W_{ij+z} = P(Z = z)/P(Z = z \mid \mathbf{V_{ij}} = v)$ for $i = 1$ to 614, $j = 1$ to J_i, where $\max(J_i) = 46$ and $z = 0$ or 1. The numerator of the weights is an estimate of the probability of being in the same calendar period as the observation; it stabilizes the weights and improves efficiency by restoring the observed distribution of Z to the weighted data (33). The denominator of the weights is an estimate of the probability of being in the same calendar period as the observation, conditional on a collection of covariates.

The collection of covariates $\mathbf{V_{ij}}$ is chosen from the set of measured covariates possibly associated with both calendar period and the endpoint based on prior information (e.g., existing literature, biology), and data exploration. Both probabilities for numerator and denominator were fit by using pooled logistic regression models (34) for the log odds of calendar period, specifically, $\text{logit}\, P(Z = z) = \gamma_0$ and $\text{logit}\, P(Z = z \mid \mathbf{V_{ij}}) = \delta_0 + \delta_1' V_{ij}$, where $\text{logit}(p) = \ln[p/(1 - p)]$, γ_0 and δ_0 are intercepts, and δ_1' is the transpose of the column vector of log odds ratios for the components of the covariate matrix $\mathbf{V_{ij}}$. $\mathbf{V_{ij}}$ included age at randomization, race/ethnicity, time since seroconversion, and the product of age at randomization and race/ethnicity. Age at randomization was centered at 35 years and was scaled by 5 years. Race/ethnicity was modeled by using a nonwhite indicator. Time since seroconversion was centered at 5 years and was fit by using a restricted cubic spline with knots at the 5th, 28th, 50th, 72nd, and 95th percentiles. This choice of knot placement minimizes the influence of the lower and upper 5% of the distribution and roughly equally splits the inner 90% (35). A description of the weights is provided in Appendix 2.

Our weighted instrumental-variable estimator of the causal rate ratio among compliers is

Table 1. Distribution of Events, Person-years, and Rates by Calendar Period and Therapy Use

Calendar Period	No. of AIDS Events	No. of Person-Years	Rate[a]
Non-HAART Therapy			
Pre-HAART	215	2,810.12	7.65
HAART	16	911.43	1.76
Total	231	3,721.55	6.21
HAART Therapy			
Pre-HAART	2	63.13	3.17
HAART	35	1,536.04	2.28
Total	37	1,599.17	2.31
Total			
Pre-HAART	217	2,873.26	7.55
HAART	51	2,447.46	2.08
Total	268	5,320.72	5.04

Abbreviations: AIDS, acquired immunodeficiency syndrome; HAART, highly active antiretroviral therapy.

[a] Rates per 100 person-years.

$$_w\beta_{\mathrm{IV}} = \frac{_w\alpha_{00} \times \left(\frac{_wD_{00}}{_wT_{00}}\right) - _w\alpha_{01} \times \left(\frac{_wD_{01}}{_wT_{01}}\right)}{_w\alpha_{11} \times \left(\frac{_wD_{11}}{_wT_{11}}\right) - _w\alpha_{10} \times \left(\frac{_wD_{10}}{_wT_{10}}\right)},$$

where $_wD_{xz} = \sum_{i=1}^{614}\sum_{j=1}^{J_i} D_{ijxz} \times W_{ijxz}$, $_wT_{xz} = \sum_{i=1}^{614}\sum_{j=1}^{J_i} T_{ijxz} \times W_{ijxz}$, and $_w\alpha_{xz} = \frac{_wT_{xz}}{_wT_{+z}}$. The leading subscript w indicates a weighted variable. We used standard formulas (36, p. 238) to compute approximate 95% confidence limits around the unweighted ITT estimates. To compute 95% confidence limits around the other estimates, we resampled the 614 men 500 times (37). We then took the standard deviation of the 500 resampled estimates on the natural logarithm (ln) scale as the standard error of the log rate ratio.

RESULTS

Table 1 gives the distribution of events, person-years, and rates by calendar period and therapy use. Overall, 268 AIDS events, 49 deaths, 90 losses to follow-up, and 207 administrative censors occurred during 5,321 person-years. In the pre-HAART calendar period, 2 of 217 (1%) events and 63 of 2,873 (2%) person-years occurred while the participant was using HAART and are therefore misclassified. In the HAART calendar period, 16 of 51 (31%) events and 911 of 2,447 (37%) person-years occurred while the participant was not using HAART and are therefore misclassified.

The odds of being in the HAART calendar period compared with the pre-HAART calendar period were 1.21-fold higher (95% CL: 1.18, 1.24) for every 5-year increase in age at seroconversion, 1.26-fold higher (95% CL: 1.08, 1.45) in the nonwhite group compared with the white group, and 5.55-fold higher (95% CL: 3.88, 7.94) for every 5-year increase in infection duration.

Table 2 gives the distribution of events and person-time by calendar period for both the unweighted and weighted data. When the pre-HAART calendar period was compared with the HAART calendar period in the unweighted data, the ITT rate ratio was 3.62 (95% CL: 2.67, 4.92) and the instrumental-variable rate ratio was 5.02 (95% CL: 3.45, 7.31). The rates in the weighted data were similar. Comparing the pre-HAART to the HAART calendar period in the weighted data showed that the ITT rate ratio was 3.44 (95% CL: 2.48, 4.77) and the instrumental-variable rate ratio was 4.95 (95% CL: 3.32, 7.37). Using 1998 as the cutoff for the HAART calendar period instead of 1996 slightly increased all rate ratio estimates: the ITT estimate became 3.74 (weighted estimate, 3.70) and the instrumental-variable estimate became 5.29 (weighted estimate, 5.13).

Figure 2 shows 3 divisions of events and person-time. We use this figure as a graphic device to illustrate the calculation of the unweighted estimators given above as β_{ITT} and β_{IV}. The first row shows the total number of events and total amount of person-time in the cohort. The second row shows the division of events and person-time by calendar period. Within calendar periods, the third row shows the division by therapy use. For those correctly classified (i.e., $x = z$), the fourth row shows how the events and person-time would have been classified assuming that calendar period z is an instrument for therapy x. The potential therapy use under calendar period $z = 0,1$ is denoted by x^z, as seen in Figure 2.

Person-time in the fourth row of Figure 2 is calculated before the number of events. Person-time for 1 calendar period is partitioned according to the conditional probability of using therapy (i.e., α_{xz}) in the other calendar period. For example, in the HAART calendar period and HAART-use arm, the 1,536.04 person-years are partitioned based on the conditional probabilities of using HAART (α_{01}) and not using HAART (α_{00}) in the pre-HAART calendar period arm: $33.75 = 0.02 \times 1,536.04$ and $1,502.29 = (1 - 0.02) \times 1,536.04$. The events for 1 calendar period are then partitioned so that the rate in the "always takers" (or "never takers") is equal to the noncomplier rate in the other calendar period. For example, the number of events in the group of HAART users from the HAART calendar period who would have used HAART had they been in the pre-HAART calendar period is chosen so that the rate in this group is equal to the rate in the group of HAART users from the pre-HAART calendar period: $1.07 = (2/63.13) \times 33.75$. The remaining events ($33.93 = 35 - 1.07$) are distributed as HAART users from the HAART calendar period who would not have used HAART had they been in the pre-HAART calendar period.

The unweighted ITT estimate comparing the pre-HAART calendar period with the HAART calendar period can be calculated by using rates from the second row of Figure 2. This calculation yields an estimate of $3.62 = (217/2,873.26)/(51/2,447.46)$, equal to that shown in Table 2. The unweighted instrumental-variable estimate can be calculated by using rates from the bottom row of Figure 2. The instrumental-variable estimate is an estimate of the complier-average causal effect. Specifically, it is the rate ratio for those in the pre-HAART calendar period who did not use HAART but would have used HAART had they been in the HAART calendar period, compared with those in the HAART calendar

Table 2. Distribution of Events and Person-time by Calendar Period

Calendar Period	No. of AIDS Events	No. of Person-Years	Rate[a]	Intent to Treat				Instrumental Variable	
				Rate Difference[a]	95% CI	Rate Ratio	95% CI	Rate Ratio	95% CI
				Unweighted[b]					
Pre-HAART	217	2,873.26	7.55	5.47	4.31, 6.62	3.62	2.67, 4.92	5.02	3.45, 7.31
HAART	51	2,447.46	2.08	0		1		1	
Total	268	5,320.72	5.04						
				Weighted[c,d]					
Pre-HAART	232.01	3,082.91	7.53	5.34	4.04, 6.64	3.44	2.48, 4.77	4.95	3.32, 7.37
HAART	51.14	2,335.93	2.19	0		1		1	
Total	283.15	5,418.84	5.23						

Abbreviations: AIDS, acquired immunodeficiency syndrome; CL, confidence limits; HAART, highly active antiretroviral therapy.

[a] Rates and rate differences per 100 person-years.

[b] Confidence intervals for unweighted intent-to treat estimates were calculated by using standard approximate formulas (36, p. 238).

[c] Confidence intervals for weighted intent-to treat estimates and instrumental-variable estimates were calculated by bootstrap (37).

[d] Covariates include age at seroconversion, race/ethnicity, time since seroconversion, and the product of age at seroconversion and race/ethnicity.

period who used HAART but would not have used HAART had they been in the pre-HAART calendar period. This calculation yields an estimate of 5.02 = (196.6/1,763.39)/(33.93/1,502.29), equal to that shown in Table 2.

DISCUSSION

In this prospective observational study, the notable amount of exposure misclassification in the HAART calendar period appeared to null-bias the rate ratio, as one would expect. Using an instrumental-variable estimator, we found that exposure to non-HAART increased the hazard of an AIDS-defining illness 5-fold compared with exposure to HAART. This result is 39% (5.02/3.62) higher than the simple calendar-period ITT result typically reported (1, 2, 21–31). To obtain these results, we extended basic instrumental-variable methods (8) from the risk difference to the rate ratio in a manner that facilitated measured-covariate adjustments based on weighting the number of events and amount of person-time by the inverse probability of calendar period given age at seroconversion, race/ethnicity, and time since seroconversion, although this latter refinement did not appreciably alter the results in our example (Table 2).

Our findings rely on the assumption that therapy use is correctly measured. Recall that antiretroviral therapy use is measured by self-report in the Multicenter AIDS Cohort
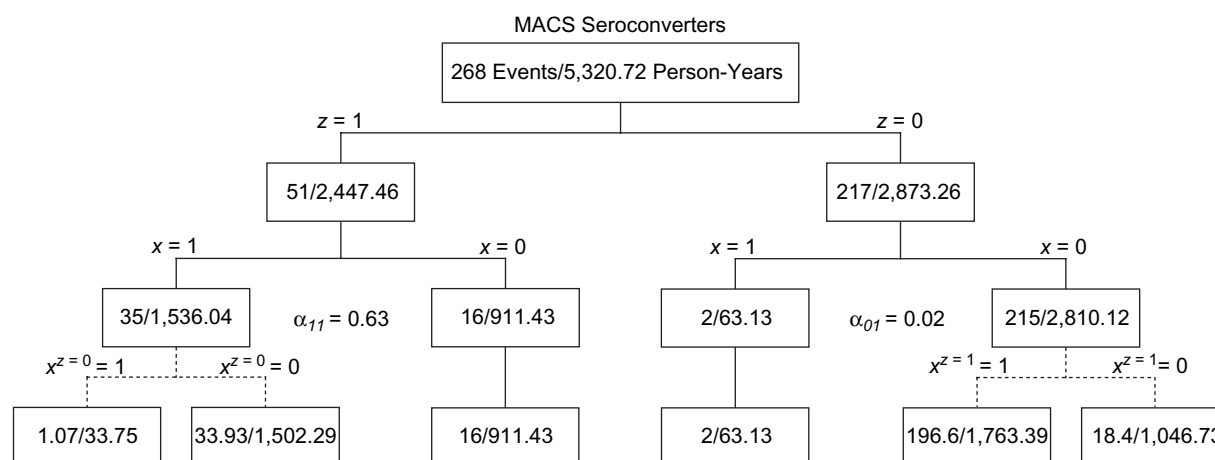
Figure 2. Tree diagram showing the division of events and person-time by 1) calendar period z, 2) therapy use x, and 3) potential therapy use x^z. Refer to the Results section of the text for more information. MACS, Multicenter AIDS Cohort Study.

Study. We would expect that violations of this assumption would result in null-biased estimates of the effect of HAART on AIDS if the misclassification is nondifferential and independent of other errors. In a sample of 68 HIV-infected patients in which self-report was compared with prescribed treatment obtained from medical record abstraction, the kappa statistic was 0.74 (95% CL: 0.55, 0.91) (38). These data suggest that self-reported HAART is a good, albeit imperfect, proxy for prescribed HAART. We further rely on the assumption that calendar period is an appropriate instrumental variable for therapy use. Conditional on controlled confounders, an instrumental variable must be 1) associated with the exposure of interest, 2) independent of the uncontrolled confounders of exposure and outcome, and 3) independent of the outcome conditional on the exposure and confounders. The first property is well documented in the current setting (1, 2, 21–31). However, properties 2 and 3 remain untestable in our observed data. The use of inverse probability of calendar period weights enables us to relax the third assumption so that the instrument (calendar period) is assumed to be independent of the outcome (AIDS) conditional on the exposure (actual therapy use) *and* controlled confounders (here, measured indications for actual therapy use).

Beyond the 3 standard assumptions, our instrumental-variable estimator, similar to that in Cuzick et al. (8), assumes exchangeability between calendar periods (17). Specifically, among the HAART users in the HAART calendar period, we assume that, had these men been observed in the pre-HAART calendar period, the same proportion of their person-time would have been on HAART as for the men observed in the pre-HAART calendar period. A parallel assumption applies to the non-HAART users in the pre-HAART calendar period. In addition, we assume that the rate among the HAART users in the HAART calendar period who would have used HAART had they been observed in the pre-HAART calendar period is equal to the rate among the HAART users in the pre-HAART calendar period. We make the same assumption for the rate among non-HAART users in the pre-HAART calendar period. This assumption could be violated if time trends are affecting comparability across the pre-HAART and HAART calendar periods such as an HIV-related non-HAART therapy introduced in the HAART calendar period that decreases the risk of AIDS.

Our findings also rely on the assumption that we have captured the full set of joint determinants of calendar period and AIDS in our model for the weights. Our final model included age at seroconversion, race/ethnicity, time since seroconversion, and the product of age at seroconversion and race/ethnicity.

While HAART was available in 1996, it was not widely used until 1998. For that reason, we also explored the impact of using 1998 as the HAART cutoff instead of 1996. The unweighted and weighted ITT estimates were 3% and 8% higher using 1998 as the cutoff, respectively. The unweighted and weighted instrumental-variable estimates were 5% and 4% higher using 1998 as the cutoff, respectively.

The present study considered only 2 calendar periods and 2 therapies. Previous studies (1, 2, 21–31) further divided our pre-HAART calendar period. Specifically, Detels et al.

(1) used 4 calendar periods: no therapy, monotherapy, combination therapy, and HAART. A logical next step would be to extend our instrumental-variable approach to 3 or more calendar periods and 3 or more therapies.

In conclusion, instrumental-variable approaches to noncompliance correction are straightforward (8, 17). Application of instrumental-variable methods to misclassification correction is equally straightforward; while well established in the measurement-error literature (39), it seems largely unrecognized in settings such as ours where an appropriate instrument is available (40). For example, calendar period has been used as an instrument in cancer research (41), and geographic location has been used as an instrument in cardiovascular research (42). Furthermore, our proposed estimator could also be adapted for use in Mendelian randomization studies (43). These methods may help to resolve discrepancies between observational and randomized evidence in a variety of biomedical fields.

Effros, John Fahey, Beth Jamieson, Otoniel Martínez-Maza, Eric N. Miller, John Oishi, Paul Satz, Harry Vinters, Dorothy Wiley, Mallory Witt, Otto Yang, Stephen Young, and Zuo Feng Zhang; *Pittsburgh*—University of Pittsburgh, Graduate School of Public Health: Charles R. Rinaldo (Principal Investigator), Lawrence Kingsley (Co-Principal Investigator), James T. Becker, Robert L. Cook, Robert W. Evans, John Mellors, Sharon Riddler, Anthony Silvestre; *Data Coordinating Center*—The Johns Hopkins Bloomberg School of Public Health: Lisa P. Jacobson (Principal Investigator), Alvaro Muñoz (Co-Principal Investigator), Stephen R. Cole, Christopher Cox, Gypsyamber D'Souza, Stephen J. Gange, Janet Schollenberger, Eric C. Seaberg, and Sol Su; *National Institutes of Health*—National Institute of Allergy and Infectious Diseases: Robin E. Huebner; *National Cancer Institute*—Geraldina Dominguez; *National Heart, Lung, and Blood Institute*—Cheryl McDonald.

For more information about this study, refer to the following website: http://www.statepi.jhsph.edu/macs/macs.html.

Conflict of interest: none declared.

REFERENCES

1. Detels R, Muñoz A, McFarlane G, et al. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Multicenter AIDS Cohort Study Investigators. *JAMA*. 1998;280(17):1497–1503.
2. Porter K, Babiker A, Bhaskaran K, et al. Determinants of survival following HIV-1 seroconversion after the introduction of HAART. *Lancet*. 2003;362(9392):1267–1274.
3. Miettinen OS. The need for randomization in the study of intended effects. *Stat Med*. 1983;2(2):267–271.
4. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med*. 1997;337(11):725–733.
5. Cameron DW, Heath-Chiozzi M, Danner S, et al. Randomised placebo-controlled trial of ritonavir in advanced HIV-1 disease. The Advanced HIV Disease Ritonavir Study Group. *Lancet*. 1998;351(9102):543–549.
6. Piantadosi S. *Clinical Trials: A Methodologic Perspective*. New York, NY: Wiley-Interscience; 1997.
7. Sommer A, Zeger SL. On estimating efficacy from clinical trials. *Stat Med*. 1991;10(1):45–52. (Erratum: *Stat Med*. 1994; 13(18):1897).
8. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med*. 1997; 16(9):1017–1029.
9. Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Commun Stat Theory Methods*. 1991;20(8):2609–2631.
10. Nagelkerke N, Fidler V, Bernsen R, et al. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Stat Med*. 2000;19(14):1849–1864.
11. Cole SR, Chu H. Effect of acyclovir on herpetic ocular recurrence using a structural nested model. *Contemp Clin Trials*. 2005;26(3):300–310.
12. Vansteelandt S, Goetghebeur E. Sense and sensitivity when correcting for observed exposures in clinical trials. *Stat Med*. 2005;24(2):191–210.
13. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*. 2000;56(3):779–788.
14. Greenland S, Lanes S, Jara M. Estimating effects from randomized trials with discontinuations: the need for intent-to-treat design and G-estimation. *Clin Trials*. 2008;5(1):5–13.
15. Cain LE, Cole SR. Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. In press.
16. Kaslow RA, Ostrow DG, Detels R, et al. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. *Am J Epidemiol*. 1987;126(2):310–318.
17. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(4):722–729.
18. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.
19. Department of Health and Human Services/Henry J Kaiser Foundation. *Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents*. 2006 (http://aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL05042006050.pdf).
20. Cole SR, Hernán MA, Robins JM, et al. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *Am J Epidemiol*. 2003;158(7):687–694.
21. Tassie JM, Grabar S, Lancar R, et al. Time to AIDS from 1992 to 1999 in HIV-1-infected subjects with known date of infection. *J Acquir Immune Defic Syndr*. 2002;30(1):81–87.
22. Tarwater PM, Mellors J, Gore ME, et al. Methods to assess population effectiveness of therapies in human immunodeficiency virus incident and prevalent cohorts. *Am J Epidemiol*. 2001;154(7):675–681.
23. Schneider MF, Gange SJ, Williams CM, et al. Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984–2004. *AIDS*. 2005;19(17): 2009–2018.
24. Palella FJ, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med*. 1998;338(13):853–860.
25. Moore RD, Chaisson RE. Natural history of HIV infection in the era of combination antiretroviral therapy. *AIDS*. 1999; 13(14):1933–1942.
26. Mocroft A, Ledergerber B, Katlama C, et al. Decline in the AIDS and death rates in the EuroSIDA study: an observational study. *Lancet*. 2003;362(9377):22–29.
27. Hoover DR, Muñoz A, He Y, et al. The effectiveness of interventions on incubation of AIDS as measured by secular increases within a population. *Stat Med*. 1994;13(19–20):2127–2139.
28. Gange SJ, Barrón Y, Greenblatt RM, et al. Effectiveness of highly active antiretroviral therapy among HIV-1 infected women. *J Epidemiol Community Health*. 2002;56(2):153–159.
29. Dorrucci M, Balducci M, Pezzotti P, et al. Temporal changes in the rate of progression to death among Italians with known date of HIV seroconversion: estimates of the population effect of treatment. Italian HIV Seroconversion Study (ISS). *J Acquir Immune Defic Syndr*. 1999;22(1):65–70.
30. Cain LE, Cole SR, Chmiel JS, et al. Effect of highly active antiretroviral therapy on multiple AIDS-defining illnesses among male HIV seroconverters. *Am J Epidemiol*. 2006; 163(4):310–315.
31. Ahdieh-Grant L, Li R, Levine AM, et al. Highly active antiretroviral therapy and cervical squamous intraepithelial

lesions in human immunodeficiency virus-positive women. *J Natl Cancer Inst.* 2004;96(14):1070–1076.

32. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep.* 1992;41(RR-17):1–19.

33. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the joint causal effect of non-randomized treatments. *J Am Stat Assoc.* 2001;96(454):440–448.

34. D'Agostino RB, Lee ML, Belanger AJ, et al. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med.* 1990; 9(12):1501–1515.

35. Harrell FE Jr. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.* New York, NY: Springer; 2001.

36. Rothman KJ, Greenland S, eds. *Modern Epidemiology.* 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998.

37. Efron B, Tibshirani R. *An Introduction to the Bootstrap.* London, United Kingdom: Chapman & Hall; 1993.

38. Cain LE. *Bridging the Gap Between Observational and Randomized Evidence: HAART and AIDS* [dissertation]. Baltimore, MD: Johns Hopkins Bloomberg School of Public Health; 2008.

39. Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models.* London, United Kingdom: Chapman & Hall/CRC; 1995.

40. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360–372.

41. Brenner H. Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis. *Lancet.* 2002;360(9340):1131–1135.

42. Stukel TA, Fisher ES, Wennberg DE, et al. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA.* 2007;297(3):278–285.

43. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22.

44. White H. *Estimation, Inference, and Specification Analysis.* Cambridge, United Kingdom: Cambridge University Press; 1993.

45. Maldonado G, Greenland S. Interpreting model coefficients when the true model form is unknown. *Epidemiology.* 1993; 4(4):310–318.

46. Gail M. Adjusting for covariates that have the same distribution in exposed and unexposed cohorts. In: Moolgavkar S, Prentice RL, eds. *Modern Statistical Methods in Chronic Disease Epidemiology.* New York, NY: Wiley; 1986:3–18.

47. Greenland S. Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology.* 1996;7(5):498–501.

48. Cole SR, Hernán MA, Margolick JB, et al. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. *Am J Epidemiol.* 2005; 162(5):471–478.

Appendix Table 1. Simulation Results Comparing the Average Causal Effect, As-Treated, Intent-to-Treat, Instrumental-Variable, and Augmented Instrumental-Variable Estimators

Sample Size, No.	Hazard and Estimator	RR	ln(RR) (MCSE)	SE	RMSE
50	Constant				
	AsT	1.49	0.40 (0.009)	0.42	0.71
	ITT	1.53	0.43 (0.009)	0.42	0.68
	IV	2.04	0.71 (0.019)	0.83	0.95
	IV2[a]	2.07	0.73 (0.019)	0.83	0.96
	Nonconstant				
	AsT	1.46	0.38 (0.009)	0.40	0.81
	ITT	1.61	0.48 (0.009)	0.39	0.71
	IV	2.23	0.80 (0.018)	0.80	0.95
	IV2	2.28	0.83 (0.019)	0.82	1.00
200	Constant				
	AsT	1.50	0.41 (0.004)	0.20	0.60
	ITT	1.54	0.43 (0.005)	0.20	0.57
	IV	1.93	0.66 (0.008)	0.35	0.48
	IV2	1.94	0.66 (0.008)	0.35	0.49
	Nonconstant				
	AsT	1.47	0.39 (0.004)	0.19	0.72
	ITT	1.61	0.48 (0.004)	0.19	0.62
	IV	2.14	0.76 (0.008)	0.36	0.55
	IV2	2.15	0.77 (0.008)	0.36	0.57
400	Constant				
	AsT	1.50	0.41 (0.003)	0.14	0.58
	ITT	1.53	0.43 (0.003)	0.14	0.56
	IV	1.89	0.64 (0.005)	0.22	0.34
	IV2	1.89	0.64 (0.005)	0.24	0.36
	Nonconstant				
	AsT	1.48	0.39 (0.003)	0.13	0.71
	ITT	1.61	0.48 (0.003)	0.13	0.60
	IV	2.09	0.74 (0.005)	0.24	0.44
	IV2	2.10	0.74 (0.005)	0.24	0.45

Abbreviations: AsT, as treated; ITT, intent to treat; IV, instrumental variable; ln(RR), arithmetic mean of the 2,000 log rate ratios; MCSE, Monte Carlo standard error; RMSE, root mean squared error of ln(RR) relative to the Kullback-Liebler Information Criterion parameter ($\approx$ln(1.82) and $\approx$ln(1.96) under a constant and nonconstant hazard, respectively) (44, 45); RR, geometric mean of the 2,000 rate ratios; SE, standard error for ln(RR).

[a] Event counts were augmented by 1/2 in each exposure group.

APPENDIX 1

Simulation of the Instrumental-Variable Estimator

To explore the amount by which the as-treated, ITT, proposed instrumental-variable, and augmented instrumental-variable estimators differ from the average causal effect, we performed 2,000 simulated draws with either a constant or nonconstant (i.e., increasing) hazard and a sample size of 50, 200, or 400. We analyzed the simulated data by using each of the methods shown in Appendix Table 1. In each scenario, N men indexed by i each had an event time $T = \min(V, W)$ generated by using a Weibull distribution, as $f(v) = \alpha \lambda v^{\alpha-1} \exp(-\lambda v^{\alpha})$, where α is the Weibull shape parameter (we use $\alpha = 1$ or 2 for constant and nonconstant hazards, respectively), $\lambda = \exp[\alpha \times \ln(2) \times x + \alpha \times \ln(4) \times u]$, X is an indicator for HAART use generated as a Bernoulli

random variable with probability $p = (1+ \exp\{-[-\ln(42.5) + \ln(4) \times u + \ln(81) \times z]\})^{-1}$, U represents the unmeasured common causes of X and T generated as a Bernoulli random variable with probability $p = 0.5$, Z is an indicator for the HAART calendar period generated as a Bernoulli random variable with probability $p = 0.5$, and random censoring time W is chosen to yield the expected 45% censored endpoints. For the average causal effect, we take the Kullback-Liebler Information Criterion (KLIC) parameter (44, 45), rather than $\alpha \times \ln(2)$, because $\alpha \times \ln(2)$ represents the covariate-conditional causal effect in nonlinear models when covariate effects are present (46, 47). The KLIC parameters (44) are $\approx \ln(1.82)$ and $\approx \ln(1.96)$ under a constant and nonconstant hazard, respectively.

As shown in Appendix Table 1, both the as-treated and ITT estimators are null biased with respect to the average causal effect but are relatively precise. The instrumental-variable estimator is less biased, albeit less precise. Augmenting the instrumental-variable estimator by adding 1/2 an event to each exposure group did not improve its performance. Combining bias and precision shows that, for the large-sample scenarios explored, the proposed instrumental-variable estimator outperforms the as-treated, ITT, and augmented instrumental-variable estimators, and, as the sample size increases, the bias in the proposed instrumental-variable estimator decreases while the biases in the as-treated and ITT estimators do not decrease.

APPENDIX 2

Description of the Weights

A box plot of the distribution of the log-standardized inverse probability of calendar period weights is displayed in Appendix Figure 1. The minimum, 1st percentile, 25th percentile, median, 75th percentile, 99th percentile, and maximum weights were 0.45, 0.67, 0.87, 0.99, 1.13, 3.46, and 32.67, respectively. The mean weight was 1.15 (standard deviation, 1.31). Truncating (48) the inverse probability of calendar period weights at the 1st and 99th percentiles yielded similar results (data not shown).
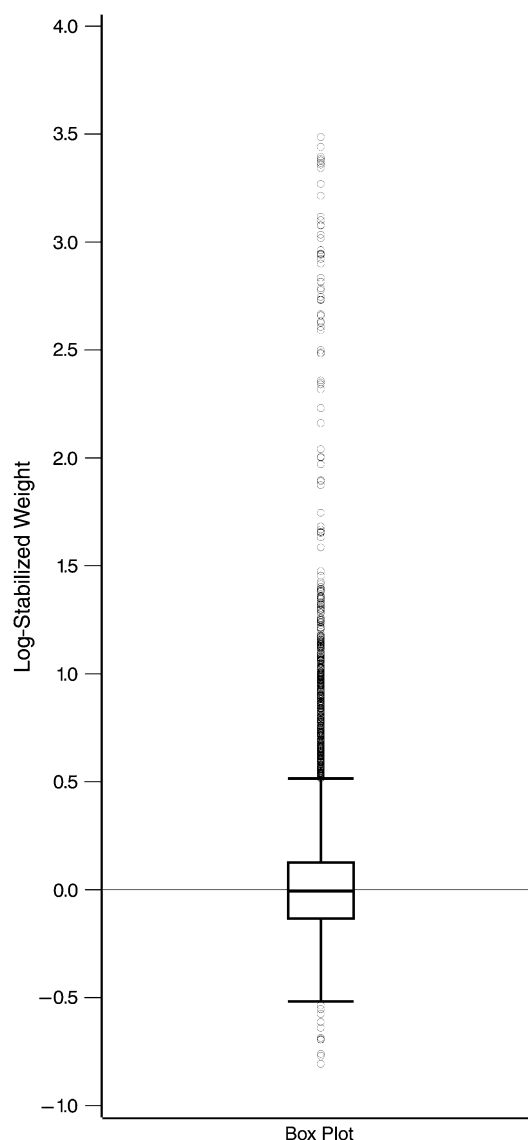
Appendix Figure 1. Distribution of inverse probability of calendar period weights.