Incident AIDS or Death After Initiation of Human Immunodeficiency Virus Treatment Regimens Including Raltegravir or Efavirenz Among Adults in the United States

Stephen R. Cole,[1] Jessie K. Edwards,[1] H. Irene Hall,[2] M. Alan Brookhart,[1] W. Christopher Mathews,[3] Richard D. Moore,[4] Heidi M. Crane,[5] Mari M. Kitahata,[5] Michael J. Mugavero,[6] Michael S. Saag,[6] and Joseph J. Eron[7]; for the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS) Investigators

[1]Department of Epidemiology, University of North Carolina, Chapel Hill; [2]Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia; [3]Department of Medicine, University of California, San Diego; [4]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland; [5]Department of Medicine, University of Washington, Seattle; [6]Department of Medicine, University of Alabama at Birmingham; and [7]Department of Medicine, University of North Carolina, Chapel Hill

Background. The long-term effectiveness of human immunodeficiency virus (HIV) treatments containing integrase inhibitors is unknown.

Methods. We use observational data from the Centers for AIDS Research Network of Integrated Clinical Systems and the Centers for Disease Control and Prevention to estimate 4-year risk of AIDS and all-cause mortality among 415 patients starting a raltegravir regimen compared to 2646 starting an efavirenz regimen (both regimens include emtricitabine and tenofovir disoproxil fumarate). We account for confounding and selection bias as well as generalizability by standardization for measured variables, and present both observational intent-to-treat and per-protocol estimates.

Results. At treatment initiation, 12% of patients were female, 36% black, 13% Hispanic; median age was 37 years, CD4 count 321 cells/μL, and viral load 4.5 $\log_{10}$ copies/mL. Two hundred thirty-five patients incurred an AIDS-defining illness or died, and 741 patients left follow-up. After accounting for measured differences, the 4-year risk was similar among those starting both regimens (ie, intent-to treat hazard ratio [HR], 0.96 [95% confidence interval {CI}, .63–1.45]; risk difference, −0.9 [95% CI, −4.5 to 2.7]), as well as among those remaining on regimens (ie, per-protocol HR, 0.95 [95% CI, .59–1.54]; risk difference, −0.5 [95% CI, −3.8 to 2.9]).

Conclusions. Raltegravir and efavirenz-based initial antiretroviral therapy have similar 4-year clinical effects. Vigilance regarding longer-term comparative effectiveness of HIV regimens using observational data is needed because large-scale experimental data are not forthcoming.

Keywords. cohort study; comparative effectiveness; HIV; raltegravir; mortality.

The comparative effectiveness of human immunodeficiency virus (HIV) treatment regimens containing the integrase inhibitor raltegravir (RAL) has been demonstrated in randomized experiments [1, 2], at least in terms of 1- to 5-year HIV virologic control. Given favorable short-term virologic control, integrase inhibitor–containing regimens have become the preferred initial treatment options in many resource-rich countries, such as the United States [3]. However, there has not yet been extensive investigation of longer-term comparative effectiveness, especially in terms of clinical endpoints such as HIV disease progression or mortality.

Using data from the National Institutes of Health–funded Centers for AIDS Research Network of Integrated Clinical Systems (CNICS) cohort, we estimate the 4-year risk of AIDS-defining illness or mortality for HIV-infected adults initiating a regimen of RAL with emtricitabine (FTC) and tenofovir disoproxil fumarate (TDF), compared to those initiating a regimen of the nonnucleoside reverse transcriptase inhibitor efavirenz (EFV) with the same backbone of FTC and TDF. We compare risk of incident AIDS or death accounting for measured differences in patient characteristics at treatment initiation, as well as measured differences between the CNICS cohort and the US population of HIV-infected adults, as estimated by the Centers for Disease Control and Prevention (CDC) [4]. Based on the favorable short-term evidence, we hypothesize that the RAL-containing regimen will provide improved 4-year outcomes compared to the comparator EFV-containing regimen.

METHODS

Study Sample

CNICS is a clinical cohort developed to support population-based HIV research in the United States [5]. CNICS includes >32 000

HIV-infected adults aged 18 or older who attended a second HIV primary care visit at one of 8 Centers for AIDS Research sites (Case Western Reserve University; Fenway Community Health Center of Harvard University; Johns Hopkins University; University of Alabama at Birmingham; University of California, San Diego; University of California, San Francisco; University of North Carolina; and University of Washington) after 1 January 1995 (or the site-specific CNICS inception date). Patient demographics, diagnoses, laboratory measurements, and medications are abstracted from point-of-care electronic medical records and compiled quarterly. Institutional review boards at each site approved study protocols. Patients provided written informed consent, or contributed administrative and/or clinical data with a waiver of written informed consent where approved by local institutional review boards.

Of the >32 000 CNICS patients, 4666 started a first-observed antiretroviral therapy regimen consisting of FTC, TDF, and either RAL or EFV while under follow-up between 12 October 2007 (US Food and Drug Administration [FDA] RAL approval date) and 1 October 2015, with all drugs started within 7 days of each other. History of antiretroviral therapy is collected at CNICS entry and patients were classified as antiretroviral naive if no evidence of prior treatment was obtained. We also excluded the 1335 patients who did not have a detectable (ie, >50 copies/mL) HIV-1 RNA viral plasma concentration measured within 180 days before to 14 days after treatment initiation, because an undetectable viral load may be indicative of unreported prior treatment. Finally, to allow control for possible bias, we excluded 270 patients (8%) due to missing data for race/ethnicity, male sex with men, injection drug use, or baseline CD4 cell count.

Endpoint Ascertainment

The outcome of interest was a first diagnosis of an AIDS-defining illness after treatment initiation, or mortality from any cause. AIDS-defining illnesses were based on the 1993 CDC clinical conditions criteria [6]. Dates of diagnoses of AIDS-defining illnesses are recorded in the electronic medical record at each site and verified before upload to CNICS. Dates of deaths are obtained from semiannual site-specific queries to the US Social Security Death Index and/or National Death Index.

Target Population

We generalize the estimated comparative effectiveness of RAL- vs EFV-containing treatment regimens in CNICS to a target population defined by all persons with HIV infection in the United States diagnosed in 2008 through 2014. The number of HIV-diagnosed persons in categories defined by race/ethnicity, sex, age group, and likely mode of transmission (ie, male sex with men, injection drug use) was provided by the CDC from national HIV surveillance data and accounts for missing information [4, 7, 8].

Statistical Methods

We compared the 4-year cumulative risk of incident AIDS or death for the RAL- and EFV-containing regimens using a treatment-decision design [9], where we restricted analyses to new users of treatment [10]. The design mimics a doubly randomized experiment where patients are randomly sampled from the target population and randomly allocated to either the RAL-containing regimen or the active comparator EFV-containing regimen [11].

Of course, the CNICS patients were neither randomly sampled from the US HIV-infected population, nor was treatment allocated at random. Therefore, we account for nonrandom sampling and treatment allocation to the extent possible using measured variables. We estimate the observational analogue to the intent-to-treat effect of starting regimens regardless of subsequent treatment changes [12], as well as the per-protocol effect of remaining on initial treatment [13].

Patients were followed from treatment initiation until first AIDS-defining illness, death, leaving observation (ie, dropouts), or administrative censoring at either 4 years after treatment initiation or 1 October 2015. To allow for possible 12-month visit intervals, patients were considered to have left observation after 18 months without a documented contact (ie, laboratory measurement or diagnosis). For the observational analogue of the per-protocol effect, patients were additionally censored at any change in treatment regimen. Two treatment regimen changes that were considered exceptions were a change from the RAL regimen to another integrase inhibitor–based regimen (including fixed-dose combination of elvitegravir, cobicistat, TDF, and FTC (FDA approved August 2012) or dolutegravir, TDF, and FTC (FDA approved August 2013), as well as a change from the EFV regimen to the nonnucleoside reverse transcriptase inhibitor regimen of rilpivirine, TDF, and FTC (FDA approved August 2011).

Risk of incident AIDS or death for each treatment regimen was estimated using the complement of the inverse probability (IP)–weighted Kaplan-Meier curve as a function of time since treatment initiation [14]. Treatment groups were compared using the hazard ratio (HR) and robust 95% confidence intervals (CIs) from Cox proportional hazards models [15], IP weighted to account for nonrandom sampling, treatment allocation, and dropout [16]. Treatment groups were also compared using the 4-year risk (probability) differences, with 95% CIs estimated from the standard deviation of 500 nonparametric bootstrap samples [17]. Time-fixed covariates included in our analysis were sex, race/ethnicity, age, history of male sex with men, history of injection drug use, a diagnosis of depression or anxiety, CD4 cell count, viral load, prevalent AIDS status, and calendar year, all measured at treatment initiation. Time-varying covariates, used to account for possible bias due to leaving observation and treatment changes, included annually updated CD4 cell count and viral load. Details of the estimation of IP weights

and the conditions under which these methods provide unbiased (technically, consistent) estimates of risk are provided in the Supplementary Appendix. We also provide the bounds [18, 19] for the treatment effect, which, while wide, apply regardless of unmeasured bias due to confounding or selection bias.

With 415 and 2646 patients in the RAL and EFV groups, respectively, and 196 endpoints in the EFV group, a 4-year risk ratio of 0.47 yields 80% statistical power for a 2-sided Wald test with type 1 error of 0.05, with expected standard error of 0.27 and expected 95% CI of .28–.80. SAS software version 9.4 (SAS Institute, Cary, North Carolina) was used for analyses.

RESULTS

Compared to the 2646 patients in the EFV group, the 415 patients in the RAL group were older, with higher CD4 cell counts, and were more likely to be female, inject drugs, and have a diagnosis of depression/anxiety or AIDS (Table 1). Compared to the 296 073 persons with HIV diagnosed between 2008 and 2014 in the United States, the 3061 patients in the CNICS sample were older and less likely to be female and black or Hispanic.

During the 4-year follow-up period, 25 of the 415 (6%) patients on RAL received a diagnosis of an AIDS-defining illness and 14 (3%) died. Of the 2646 patients receiving EFV, 144 (5%) were diagnosed with an AIDS-defining illness and 52 (2%) died. Dropout was similar across the treatment groups with 23% (94/415) of the RAL group and 24% (647/2646) of the EFV group leaving follow-up event-free before 4 years. Details of patient follow-up are provided in Supplementary Figure 1.

The crude intent-to-treat 4-year risk of incident AIDS or death from any cause, as a function of time from treatment initiation, was 9.9% among those who started the RAL regimen and 8.1% among those who started the EFV regimen, as shown in the upper left panel of Figure 1. The crude intent-to-treat 4-year risk difference was 1.8 (95% CI, −1.4 to 5.0), and the hazard ratio was 1.32 (95% CI, .94–1.87; Table 2). The bounds for all possible values of the 4-year risk difference under no assumptions about confounding or selection bias were −35% under the best-case scenario for RAL, and 65% under the worst-case scenario for RAL (Supplementary Figure 2).

After accounting for nonrandom sampling, treatment allocation, and dropout, the IP-weighted intent-to-treat 4-year risk of incident AIDS or death from any cause was 8.4% among those who started the RAL regimen and 9.3% for those who started the EFV regimen, as shown in the upper right panel of Figure 1. The IP-weighted intent-to-treat 4-year risk difference was −0.9 (95% CI, −4.5 to 2.7), and the hazard ratio was 0.96 (95% CI, .63–1.45; Table 2). Changes in the intent-to-treat hazard ratio due to each of nonrandom sampling, treatment allocation, and dropout are detailed in Supplementary Table 1.

Eighteen percent (76/413) of the RAL group, and 8% (201/2646) of the EFV group, had a protocol-allowed treatment change before incurring an outcome or completing the study. Twenty percent (84/415) of the RAL group, and 28% (733/2646) of the EFV group, had a protocol-ending treatment change before incurring an outcome or completing the study. In the observed data, the crude per-protocol 4-year

---

Table 1. Characteristics at Treatment Initiation Among 3061 Human Immunodeficiency Virus (HIV)–Infected Adults in Care Between 12 October 2007 and 1 October 2015, and Among US Target Population Who Received a Diagnosis of HIV Between 2008 and 2014

Characteristic	Raltegravir Regimen[a] (n = 415)	Efavirenz Regimen[a] (n = 2646)	Overall (n = 3061)	US Target Population (n = 296 073)
Female sex	19 (78)	11 (302)	12 (380)	21 (63 140)
Black race	33 (135)	36 (958)	36 (1093)	45 (134 646)
Hispanic ethnicity	9 (39)	13 (344)	13 (383)	22 (63 830)
Age, y, median (IQR)	40 (32–48)	37 (29–46)	37 (29–46)	NA
Age, y	…	…	…	…
18–24	6 (25)	11 (296)	10 (321)	20 (59 274)
25–34	27 (114)	31 (833)	31 (947)	29 (84 781)
35–44	30 (124)	29 (776)	29 (900)	23 (68 708)
45–54	26 (107)	22 (582)	23 (689)	19 (55 800)
≥55	11 (45)	6 (159)	7 (204)	9 (27 510)
Male sex with men	64 (267)	72 (1907)	71 (2174)	65 (192 742)
Injection drug use	16 (67)	11 (300)	12 (367)	8 (23 134)
AIDS diagnosis[b]	15 (61)	12 (310)	12 (371)	NA
Depression or anxiety	25 (105)	21 (553)	22 (658)	NA
CD4 count, cells/µL, median (IQR)	347 (181–509)	319 (179–459)	321 (180–466)	NA
Viral load, copies/mL, median (IQR)	4.4 (3.5–5.0)	4.5 (3.8–5.1)	4.5 (3.7–5.1)	NA
Calendar year, median (IQR)	2011 (2010–2013)	2010 (2009–2012)	2010 (2009–2012)	2011 (2009–2013)

Data are presented as percentage (No.) unless otherwise indicated.

Abbreviations: IQR, interquartile range; NA, not available; US, United States.

[a]Both regimens included the same backbone of emtricitabine and tenofovir disoproxil fumarate.

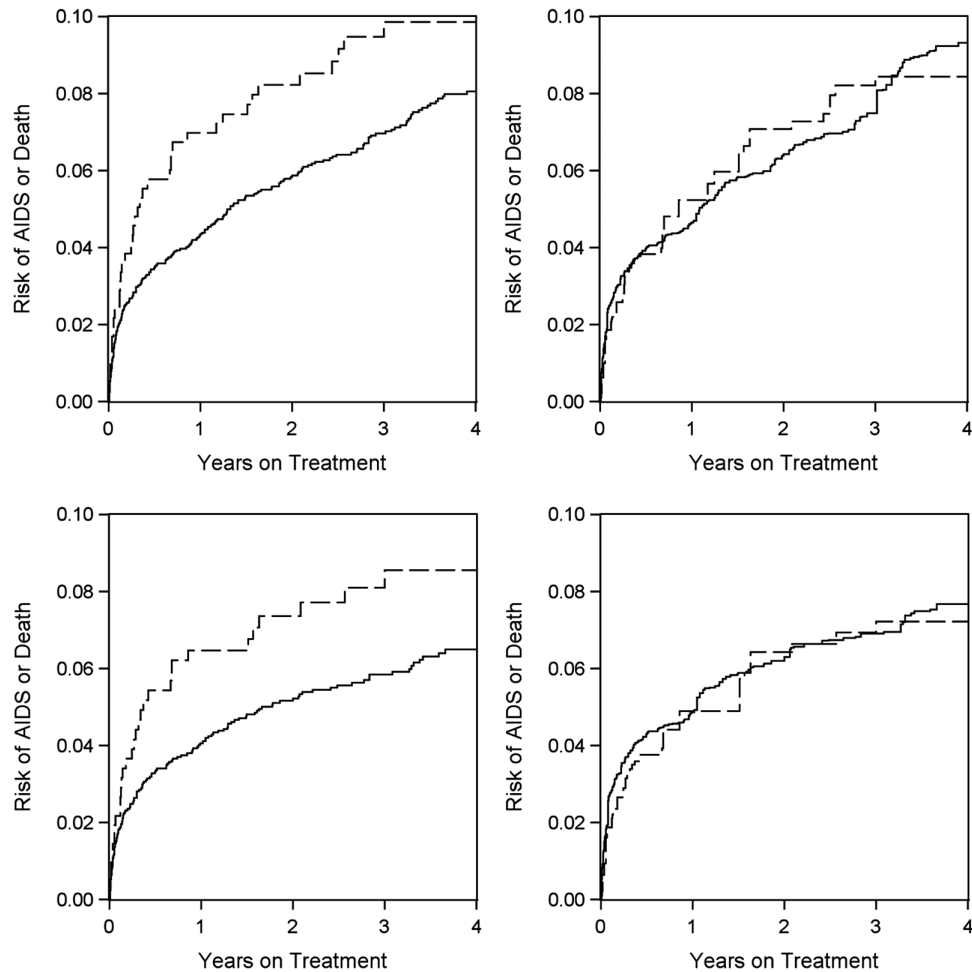[b]Prevalent AIDS diagnosis at treatment initiation.

Figure 1. Risk of incident AIDS or death from any cause over 4 years from treatment initiation among 3061 human immunodeficiency virus–infected adults in care between 12 October 2007 and 1 October 2015. Solid line represents the efavirenz regimen, and dashed line represents the raltegravir regimen. Upper panels represent intent-to-treat analyses with the left panel depicting the crude results, and right panel the inverse probability (IP)–weighted results (accounting for nonrandom sampling, treatment allocation, and dropout). Lower panels represent per-protocol analyses with the left panel again representing crude results, and the right panel the IP-weighted results (accounting for nonrandom sampling, treatment allocation, dropout, and treatment changes).

risk of incident AIDS or death from any cause, as a function of time from treatment initiation, was 8.6% among those in the RAL group and 6.5% for those in the EFV group, as shown in the lower left panel of Figure 1. The crude per-protocol 4-year risk difference was 2.1 (95% CI, −1.0 to 5.1), and the HR was 1.41 (95% CI, .96–2.07; Table 2). After accounting for nonrandom sampling, treatment allocation, dropout, and treatment changes, the IP-weighted per-protocol 4-year risk of incident AIDS or death from any cause was 7.2% among those who started the RAL regimen and 7.7% for those who started the EFV regimen, as shown in the lower right panel of Figure 1. The IP-weighted per-protocol 4-year risk difference was −0.5 (95% CI, −3.8 to 2.9), and the HR was 0.95 (95% CI, .59–1.54; Table 2). Changes in the per-protocol HR due to each of nonrandom sampling, treatment allocation, dropout, and treatment changes are detailed in Supplementary Table 1.

DISCUSSION

In a large multisite clinical cohort, we first observed apparently higher risk of 4-year HIV disease progression among patients starting and remaining on an antiretroviral therapy combination including RAL, compared to those starting and remaining on a combination including EFV with the same backbone using a per-protocol analysis. However, the apparent difference in incident AIDS or all-cause death was largely eliminated upon jointly accounting for nonrandom sampling, treatment allocation, and dropout. Crude and adjusted results were similar when we ignored treatment changes using an observational equivalent of intent-to-treat analysis. Findings indicate a similar 4-year clinical effect of RAL or EFV initial antiretroviral therapy, with the same backbone of FTC and TDF.

Existing randomized evidence has concentrated on relatively short-term virologic endpoints and lack the statistical power to evaluate clinical endpoints, including mortality. Efavirenz and RAL

Table 2. Risk of Incident AIDS or Death From Any Cause Over 4 Years From Treatment Initiation Among 3061 Human Immunodeficiency Virus–Infected Adults in Care Between 12 October 2007 and 1 October 2015

Analysis	Patients, No.	Person-years, No.	Outcomes, No.	4-Year Risk	4-Year Risk Difference (95% CI)	Hazard Ratio (95% CI)
Intent-to-treat analyses						
Crude						
Efavirenz[a]	2646	8604.9	196	8.1	0	1
Raltegravir[a]	415	1232.6	39	9.9	1.8 (−1.4 to 5.0)	1.32 (.94–1.87)
IP-weighted[b]						
Efavirenz[a]	…	…	…	9.3	0	1
Raltegravir[a]	…	…	…	8.4	−0.9 (−4.5 to 2.7)	0.96 (.63–1.45)
Per-protocol analyses						
Crude						
Efavirenz[a]	2646	6858.8	143	6.5	0	1
Raltegravir[a]	415	1072.9	32	8.6	2.1 (−1.0 to 5.1)	1.41 (.96–2.07)
IP-weighted[b]						
Efavirenz[a]	…	…	…	7.7	0	1
Raltegravir[a]	…	…	…	7.2	−0.5 (−3.8 to 2.9)	0.95 (.59–1.54)

Abbreviations: CI, confidence interval; IP, inverse probability.

[a]Both regimens included the same backbone of emtricitabine and tenofovir disoproxil fumarate.

[b]Results from IP-weighted Kaplan-Meier and Cox regression estimates with IP weights for sampling, treatment regimen, and dropout (and regimen change for per-protocol analyses).

have only been compared once in a large randomized clinical trial of treatment-naive patients [1, 20]. In that trial, RAL or EFV was paired with FTC and tenofovir: virologic suppression rates were similar at 48 and 96 weeks. After 5 years of follow-up, virologic suppression was greater in the RAL arm, driven predominantly by discontinuation in the EFV arm [21]. Serious clinical events and deaths were uncommon and similar between arms. Other integrase inhibitors have been compared to EFV in randomized trials in treatment-naive patients and results were similar to those seen with RAL. One trial demonstrated superiority of the integrase inhibitor in maintaining virologic suppression over 96 weeks, again driven by discontinuations in an EFV arm [22, 23], and another trial demonstrated similar viral suppression outcomes between the integrase inhibitor arm and an EFV arm [24–26].

Our results are subject to limitations. Foremost, patient sampling, treatment allocation, dropout, and treatment changes were not randomized. It does not appear that sicker individuals (defined by lower CD4 cell count, higher viral load, and prevalent AIDS diagnosis) were more likely to receive the RAL regimen. However, physicians may be channeling patients to RAL for reasons not captured in the set of adjustment variables. For example, RAL may be prescribed preferentially for those with cardiovascular or liver disease [27]. Moreover, patients expected to be less compliant may be prescribed RAL preferentially, thereby causing a weakened protective effect of RAL vs EFV. Dropout and treatment changes were relatively balanced across treatment groups. While we accounted for differences due to measured variables, there might have been unmeasured causes in common with AIDS or death, which could bias our results. Second, our results are relatively imprecise due to the small number of outcomes in the RAL group, which is a result of a combination of recent uptake and treatment effectiveness. Indeed, we allow patients to enter follow-up beginning 12 October 2007 when RAL was FDA approved, but few patients started first-line RAL before 2009 when it was recommended in the US Department of Health and Human Services guidelines. In particular, the small size of the RAL group precludes our ability to explore how risk differs for subgroups (eg, age). Moreover, the improved effectiveness of antiretroviral therapies and the small number of treatment-naive patients with long-term integrase inhibitor experience make differences in risk especially difficult to detect in randomized or nonrandomized studies. Third, our results might be biased due to measurement errors. However, data from both the CNICS and the CDC are produced under strict quality guidelines. Moreover, key variables used here (including demographic characteristics, treatment regimen, AIDS diagnoses, and deaths) are carefully collected and reviewed or adjudicated and likely measured with only negligible error. We did not provide details regarding cause of death; quantifying cause of death is difficult generally, and particularly so in the modern era of HIV as relatively sharply defined AIDS-related deaths recede and a large set of heterogeneous causes emerges. Fourth, we assume correct specification of the models used to estimate the inverse probability weights.

The data employed here are among the highest quality and quantity available in the United States from HIV observational cohorts, and experimental evidence is not forthcoming. Observational cohort studies collecting comprehensive longitudinal data provide a valuable source of information supplementing estimates from randomized trials. In the absence of data from randomized trials, rich prospective observational data coupled with logically principled quantitative methods

provide the best available evidence for assessment of comparative effectiveness. In our study, reweighting of the observed data [28, 29] accounted for measured imbalances in treatment allocation and dropout (and treatment stops) and so removed confounding and selection bias associated with observed variables. This reweighting provides an internally valid estimate of the treatment effect, assuming we measured (and correctly modeled) a correct set of variables. The data were further reweighted to account for measured imbalances between the CNICS sample and the CDC population, and therefore removed sampling bias associated with observed variables. This accounting provides an externally valid estimate of the treatment effect, assuming we measured (and correctly modeled) a correct set of variables. Although there is some experimental evidence for a short and moderate-term advantage of integrase inhibitor regimens on viremia, our analysis indicates a similar clinical effect to 4 years. Continuing assessment of the longer-term comparative effectiveness of these regimens using observational data is needed.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

References

1. Lennox JL, DeJesus E, Lazzarin A, et al; STARTMRK investigators. Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naive patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. Lancet **2009**; 374:796–806.

2. Eron JJ, Cooper DA, Steigbigel RT, et al; BENCHMRK Study Teams. Efficacy and safety of raltegravir for treatment of HIV for 5 years in the BENCHMRK studies: final results of two randomised, placebo-controlled trials. Lancet Infect Dis **2013**; 13:587–96.

3. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Atlanta, GA: Centers for Disease Control and Prevention, **2016**.

4. Centers for Disease Control and Prevention. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data—United States and 6 dependent areas, 2013. HIV Surveillance Supplemental Report 20 (No. 2). Atlanta, GA: CDC, **2015**.

5. 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.

6. 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 Recomm Rep **1992**; 41(RR-17):1–19.

7. Centers for Disease Control and Prevention. HIV surveillance report 2011. Atlanta, GA: CDC, **2013**.

8. Centers for Disease Control and Prevention. Diagnoses of HIV infection in the United States and dependent areas 2014. HIV Surveillance Report, vol 26. Atlanta, GA: CDC, **2015**.

9. Brookhart MA. Counterpoint: the treatment decision design. Am J Epidemiol **2015**; 182:840–5.

10. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol **2003**; 158:915–20.

11. 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.

12. Cole SR, Hernán MA, Margolick JB, Cohen MH, Robins JM. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. Am J Epidemiol **2005**; 162:471–8.

13. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. Clin Trials **2012**; 9:48–55.

14. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Statist Assoc **1958**; 53:457–81.

15. Cox DR. Regression models and life tables. J R Statist Soc (B) **1972**; 34:187–220.

16. Buchanan AL, Hudgens MG, Cole SR, Lau B, Adimora AA; Women's Interagency HIV Study. Worth the weight: using inverse probability weighted Cox models in AIDS research. AIDS Res Hum Retroviruses **2014**; 30:1170–7.

17. Efron B, Tibshirani R. An introduction to the bootstrap. London: Chapman Hall, **1993**.

18. Robins JM. The analysis of randomized and nonrandomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest L, Freeman H, Mulley A. Health service research methodology: a focus on AIDS. Washington, DC: US Public Health Service, **1989**:113–59.

19. Cole SR, Hudgens MG, Edwards JK. A fundamental equivalence between randomized experiments and observational studies. Epidemiologic Methods **2016**; 5:113–7.

20. Lennox JL, Dejesus E, Berger DS, et al; STARTMRK Investigators. Raltegravir versus efavirenz regimens in treatment-naive HIV-1-infected patients: 96-week efficacy, durability, subgroup, safety, and metabolic analyses. J Acquir Immune Defic Syndr **2010**; 55:39–48.

21. Rockstroh JK, DeJesus E, Lennox JL, et al; STARTMRK Investigators. Durable efficacy and safety of raltegravir versus efavirenz when combined with tenofovir/emtricitabine in treatment-naive HIV-1-infected patients: final 5-year results from STARTMRK. J Acquir Immune Defic Syndr **2013**; 63:77–85.

22. Walmsley SL, Antela A, Clumeck N, et al; SINGLE Investigators. Dolutegravir plus abacavir-lamivudine for the treatment of HIV-1 infection. N Engl J Med **2013**; 369:1807–18.

23. Walmsley S, Baumgarten A, Berenguer J, et al. Brief report: dolutegravir plus abacavir/lamivudine for the treatment of HIV-1 infection in antiretroviral therapy-naive patients: week 96 and week 144 results from the SINGLE randomized clinical trial. J Acquir Immune Defic Syndr **2015**; 70:515–9.

24. Sax PE, DeJesus E, Mills A, et al; GS-US-236-0102 study team. Co-formulated elvitegravir, cobicistat, emtricitabine, and tenofovir versus co-formulated efavirenz, emtricitabine, and tenofovir for initial treatment of HIV-1 infection: a randomised, double-blind, phase 3 trial, analysis of results after 48 weeks. Lancet **2012**; 379:2439–48.

25. Zolopa A, Sax PE, DeJesus E, et al; GS-US-236-0102 Study Team. A randomized double-blind comparison of coformulated elvitegravir/cobicistat/emtricitabine/tenofovir disoproxil fumarate versus efavirenz/emtricitabine/tenofovir disoproxil fumarate for initial treatment of HIV-1 infection: analysis of week 96 results. J Acquir Immune Defic Syndr **2013**; 63:96–100.

26. Wohl DA, Cohen C, Gallant JE, et al; GS-US-236-0102 Study Team. A randomized, double-blind comparison of single-tablet regimen elvitegravir/cobicistat/emtricitabine/tenofovir DF versus single-tablet regimen efavirenz/emtricitabine/tenofovir DF for initial treatment of HIV-1 infection: analysis of week 144 results. J Acquir Immune Defic Syndr **2014**; 65:e118–20.

27. Saag MS, Westfall AO, Cole SR, et al. Brief report: factors associated with the selection of initial antiretroviral therapy from 2009 to 2012. J Acquir Immune Defic Syndr **2017**; 74:60–4.

28. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology **2000**; 11:550–60.

29. Robins JM. Association, causation and marginal structural models. Synthese **1999**; 121:151–79.