American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2023.

Vol. 192, No. 8
https://doi.org/10.1093/aje/kwad057
Advance Access publication:
March 15, 2023

Original Contribution

Optimizing Treatment for Human Immunodeficiency Virus to Improve Clinical Outcomes Using Precision Medicine

Michael Jetsupphasuk*, Michael G. Hudgens, Haidong Lu, Stephen R. Cole, Jessie K. Edwards, Adaora A. Adimora, Keri N. Althoff, Michael J. Silverberg, Peter F. Rebeiro, Viviane D. Lima, Vincent C. Marconi, Timothy R. Sterling, Michael A. Horberg, M. John Gill, Mari M. Kitahata, Richard D. Moore, Raynell Lang, Kelly Gebo, Charles Rabkin, and Joseph J. Eron

* Correspondence to Michael Jetsupphasuk, Department of Biostatistics, UNC Gillings School of Global Public Health, Chapel Hill, NC 27599 (e-mail: jetsupphasuk@unc.edu).

Initially submitted July 21, 2022; accepted for publication March 9, 2023.

In first-line antiretroviral therapy (ART) for human immunodeficiency virus (HIV) treatment, some subgroups of patients may respond better to an efavirenz-based regimen than an integrase strand transfer inhibitor (InSTI)-based regimen, or vice versa, due to patient characteristics modifying treatment effects. Using data based on nearly 16,000 patients from the North American AIDS Cohort Collaboration on Research and Design from 2009–2016, statistical methods for precision medicine were employed to estimate an optimal treatment rule that minimizes the 5-year risk of the composite outcome of acquired immune deficiency syndrome (AIDS)-defining illnesses, serious non-AIDS events, and all-cause mortality. The treatment rules considered were functions that recommend either an efavirenz- or InSTI-based regimen conditional on baseline patient characteristics such as demographic information, laboratory results, and health history. The estimated 5-year risk under the estimated optimal treatment rule was 10.0% (95% confidence interval (CI): 8.6, 11.3), corresponding to an absolute risk reduction of 2.3% (95% CI: 0.9, 3.8) when compared with recommending an efavirenz-based regimen for all patients and 2.6% (95% CI: 1.0, 4.2) when compared with recommending an InSTI-based regimen for all. Tailoring ART to individual patient characteristics may reduce 5-year risk of the composite outcome compared with assigning all patients the same drug regimen.

antiretroviral therapy; efavirenz; human immunodeficiency virus; integrase strand transfer inhibitors; optimal treatment rule; precision medicine

Abbreviations: AIDS, acquired immune deficiency syndrome; ART, antiretroviral therapy; CI, confidence interval; HIV, human immunodeficiency virus; InSTI, integrase strand transfer inhibitor; NA-ACCORD, North American AIDS Cohort Collaboration on Research and Design; SAGA, smoothed and augmented genetic algorithm.

First-line antiretroviral therapy (ART) for human immunodefficiency virus (HIV) has shifted from nonnucleoside reverse transcriptase inhibitor (NNRTI)-containing regimens to integrase strand transfer inhibitors (InSTI)-containing regimens, based on evidence from clinical trials that demonstrated InSTIs led to more rapid viral suppression, fewer side effects, and less drug resistance (1–3). When examining long-term patient outcomes, a recent observational study found no appreciable difference between regimens based on efavirenz (an NNRTI) and InSTI (4). Other observational studies have been limited by insufficient sample size or follow-up, precluding precise estimation of treatment effects for long-term outcomes (5, 6). However, there may exist variability in treatment effects, whereby patients with certain characteristics may benefit more from one treatment over another. Under sufficient treatment effect heterogeneity, there may exist a treatment rule (i.e., a function that recommends treatment based on patient characteristics such as demographic information, laboratory results, or health history) that improves patient outcomes compared with assigning all patients the same treatment.

Certain statistical methods for precision medicine attempt to find an optimal treatment rule that recommends different treatments to subgroups of patients using data based on patient characteristics in order to best improve patient outcomes (7, 8). In clinical practice, medical providers consider individuals on a case-by-case basis and recommend treatments accordingly. However, this clinician-level process is difficult to evaluate and scale. Precision medicine methods, on the other hand, can estimate treatment rules that, when followed, may improve average patient outcomes on a population basis, under some requisite assumptions.

In this study, statistical methods for precision medicine were applied to an observational data set of nearly 16,000 adult patients diagnosed with HIV to estimate an optimal treatment rule that minimizes 5-year risk of the composite outcome of acquired immune deficiency syndrome (AIDS)-defining illnesses, serious non-AIDS events, and all-cause mortality. Two treatment groups were considered: those who initiated an efavirenz-based drug regimen and those who initiated an InSTI-based drug regimen.

METHODS

Study population and eligibility criteria

This study used data from the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD), a collaboration of prospective cohort studies that collects data on adults living with HIV. The analysis data set included 15,993 patients who were HIV-seropositive, ART-naive, and at least 18 years old, and who had initiated an efavirenz- or InSTI-based regimen between July 2009 and December 2016. The class of InSTIs included raltegravir (RAL), dolutegravir (DTG), and elvitegravir-cobicistat (EVG/COBI). The backbones for both regimens consisted of emtricitabine and one of either tenofovir disoproxil fumarate or tenofovir alafenamide. Patients with a history of acute myocardial infarction or stroke, end-stage renal disease (ESRD), or end-stage liver disease (ESLD) were excluded since these conditions are long-term or recurring and were part of the composite outcome of interest.

The study sample, as well as the outcome definition and covariates (described below), were similar to a previous analysis of the NA-ACCORD (4). A brief comparison of the previous study with the analysis presented here is included in Web Appendix 1 (available at https://doi.org/10.1093/aje/kwad057).

Outcome

The composite outcome was defined analogously to other recent studies on ART and included the first occurrence of an AIDS-defining illness, serious non-AIDS event, or all-cause death (4, 9, 10). An AIDS-defining illness was defined according to the criteria published by the Centers for Disease Control and Prevention in 1993 (11), and serious non-AIDS events consisted of acute myocardial infarction or stroke, ESRD, and ESLD. ESRD was defined as an estimated glomerular filtration rate consistently less than 30 mL/minute/1.73 m^2 for at least 3 months, and ESLD was defined to be 2 fibrosis-4 scores greater than 3.25, more than 6 months apart (4).

Covariates

Three sets of covariates were used in this analysis: one set for baseline confounding adjustment, one set for differential loss to follow-up adjustment, and one set for treatment rule inputs. Covariates considered as potential confounders included baseline measurements of: CD4 T-cell count, HIV viral load, age, body mass index, calendar year at initiation, and indicators of female sex, Black race, Hispanic ethnicity, men who have sex with men, injection drug use, risky heterosexual sex, history of any clinical AIDS diagnosis, hepatitis B infection (a positive antigen test, a positive surface antigen test, or a positive DNA test result), hepatitis C infection (presence of hepatitis C genotype test, detectable RNA, or a positive antibody test), depression diagnosis, anxiety diagnosis, diabetes mellitus (glycosylated hemoglobin at least 6.5%, medication specifically for diabetes, or a diagnosis with a diabetes-related medication), hypertension (clinical diagnosis and prescription of antihypertensive medication), elevated total cholesterol (at least 240 mg/dL), and statin prescription (4). Since the focus of this study is on the effect of ART initiation, only baseline covariates were considered as potential confounders. The set of covariates used to correct for differential loss to follow-up included the baseline covariates and time-varying measurements of CD4 T-cell count, HIV viral load, diabetes mellitus, depression, anxiety, hypertension, elevated total cholesterol, and statin prescription (4). Since censoring occurred at various times after baseline, time-varying measurements were included to adjust for differential loss to follow-up. Covariates considered possible effect modifiers were inputs for the treatment rules and included antidepressant use (see Web Table 1 for the list of antidepressants) at baseline and all the other baseline covariates except calendar year at initiation, since the goal of developing an optimal treatment rule is to inform treatment selection prospectively. Measurements of age, sex, race, ethnicity, and behavior were collected at enrollment. Baseline measurements of CD4 T-cell count and HIV viral load were collected from 90 days before to 7 days after ART initiation. All other baseline covariates were measured at ART initiation.

Missing covariates were imputed using a random forest–based approach (12). The analysis was also performed after using multiple imputations by chained equations to handle missing data, and results were similar (not shown). Web Table 2 provides a summary of missingness for each covariate. CD4 T-cell count and HIV viral load were missing for 22% and 25% of participants, respectively; there was a negligible amount of missingness for other covariates. Since measurements for CD4 T-cell count and HIV viral load occurred within a window around ART initiation, it is plausible that missingness was completely at random for patients who did not have a clinical appointment in that window. For patients who did have appointments but missed them, it is plausible that the missing data was missing at random conditional on controlling for measured medical conditions and illnesses.

Statistical methods for precision medicine

Ten different statistical methods for precision medicine were used to estimate the treatment rule that minimizes the 5-year composite disease risk: penalized Cox regression (with ridge, lasso, and elastic net penalties) (13, 14), causal survival forests (15), a method using a genetic algorithm (with 4 variations) (16), outcome-weighted learning (17), and residual-weighted learning (18). An evaluation procedure, described below, was used to select the precision medicine method that minimized estimated risk. For methods that produced similarly low risk estimates, confidence interval width was used as a secondary criterion. The selected method was then used to estimate the optimal treatment rule, which was encoded in an Excel (Microsoft, Redmond, Washington) file, provided in Web Appendix 2.

The penalized Cox regression and causal survival-forests methods model the outcome to estimate the optimal treatment rule. These outcome regression methods rely on correct model specification but may be statistically more efficient than the other methods considered if the model specification is correct. The genetic algorithm method uses computational techniques to directly optimize a risk estimator with respect to a class of treatment rules. Four variations of the genetic algorithm were employed. The smoothed and augmented genetic algorithm (SAGA) method employs a risk estimator with a smoothed version of the treatment rule and with augmentation using an outcome regression model. The 3 other variations include only smoothing, only augmentation, or neither. Outcome-weighted learning estimates optimal treatment rules by transforming the optimization problem to one solvable using support vector machines, and residual-weighted learning extends outcome-weighted learning to incorporate an outcome regression model.

The outcome of optimization for the penalized Cox regression and genetic algorithm methods was the 5-year risk of the composite outcome, while causal survival forests and outcome-/residual-weighted learning targeted the related measure of restricted mean survival time with respect to the same composite indicator. All methods under consideration, except the causal survival forests, require a class of treatment rules to be prespecified, and the class of linear treatment rules was chosen.

See Web Appendix 3 for detailed descriptions of the implemented statistical methods for precision medicine. SAS, version 9.4 (SAS Institute, Inc., Cary, North Carolina), or R, version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria), was used for all data cleaning or analysis.

Evaluation

The 5-year composite disease risk was estimated using an inverse probability of treatment and censoring–weighted Kaplan-Meier (IPW-KM) estimator that corrects for baseline confounding and differential loss to follow-up (16). The IPW-KM estimator estimates risk as a function of treatment rules, and the mathematical expression of the estimator can be found in Web Appendix 3. The risk under the estimated optimal treatment rule was estimated and compared with the estimated risks under the specific treatment rules where all patients are assigned to an efavirenz-based regimen, all patients are assigned to an InSTI-based regimen, and all patients are assigned according to their observed treatment decisions. In addition to estimating risks for the composite outcome, cumulative incidence functions were estimated to assess the performance of the estimated optimal treatment rule on each component of the composite outcome (19).

The optimal treatment rule is the treatment rule resulting in the lowest risk of the 5-year composite outcome. For each of the 10 precision medicine methods, the optimal treatment rule was estimated using a cross-validation procedure to avoid overfitting. The sample splitting procedure separated data used to estimate treatment rules and risk (20). Missing data imputation and models for correcting baseline confounding and differential loss to follow-up were computed before the cross-validation procedure. The procedure was used for all methods, with the exception that optimization starting values for the genetic algorithm methods were based on the whole data set.

The cross-validation procedure split the data into 10 groups, where one group was considered the test set and the remaining 9 groups were considered the training set. A treatment rule was estimated using each training set and then used to output treatment assignments in the corresponding test sample. The process was then repeated with a different group as the test set until each of the 10 groups was considered a test set exactly once. The cross-validated risk estimate was then computed by evaluating the IPW-KM estimator under the test set treatment assignments. Standard errors were estimated using 500 bootstraps (without bias correction) where each bootstrap resampled the data randomly with replacement. Only estimation of treatment rules and risk were included in each bootstrap resample; missing data imputation was performed prior to the bootstrap procedure. A sensitivity analysis was performed to account for uncertainty of the missing data imputation; see Web Appendix 4 and Web Table 3 for more details. In a simulation study (details provided in Web Appendix 5 and Web Table 4; see data availability statement for code), Wald confidence intervals (CIs) based on bootstrap standard error estimates were shown to have nominal to slightly above nominal coverage at the 95% level. Given the computational burden of estimating standard errors and the analysis objective of minimizing 5-year risk, standard errors were estimated for only the efavirenz risk, InSTI risk, observed treatment risk, and the smallest risk estimates produced by a precision medicine treatment rule.

RESULTS

There were 10,169 (63.6%) participants who initiated an efavirenz-based regimen and 5,824 (36.4%) who initiated an InSTI-based regimen. The percentage of participants who initiated each drug regimen varied over calendar time as guidelines evolved. In 2009, only 9.2% of patients initiated an InSTI-based regimen, but by 2016, 90.9% of patients initiated an InSTI-based regimen. Summary statistics of nonimputed patient characteristics for each treatment group are given in Table 1. Of the 5,824 patients assigned to InSTI, 1,840 (31.6%) initiated raltegravir, 623 (10.7%)

Table 1. Characteristics of 15,993 Adults Living With HIV, Who Initiated an Integrase Strand Transfer Inhibitor–Based or Efavirenz-Based Antiretroviral Therapy Regimen in the North American AIDS Cohort Collaboration on Research and Design, 2009–2016

Characteristic	Participants According to Observed Regimen					
	InSTI-Based[a] (*n* = 5,824)		EFV-Based[a] (*n* = 10,169)		Overall (*n* = 15,993)	
	No.	%	No.	%	No.	%
Age, years[b]	37 (28.0–48.0)		41.0 (31.0–50.0)		40.0 (30.0–50.0)	
Female sex	894	15.3	1,101	10.8	1,995	12.5
Black race	2,351	40.4	4,611	45.3	6,962	43.5
Hispanic ethnicity	722	12.4	1,350	13.3	2,072	13.0
Body mass index[b,c]	25.1 (22.3–28.7)		25.1 (22.3–28.6)		25.1 (22.3–28.6)	
Injection drug use	566	9.7	1,033	10.2	1,599	10.0
Male-to-male sexual contact	3,209	55.1	4,523	44.5	7,732	48.3
Heterosexual behavior	1,349	23.2	2,030	20.0	3,379	21.1
Previous AIDS diagnosis	480	8.2	735	7.2	1,215	7.6
Hepatitis B	214	3.7	426	4.2	640	4.0
Hepatitis C	564	9.7	1,123	11.0	1,687	10.5
Previous depression diagnosis	859	14.7	1,038	10.2	1,897	11.9
Previous anxiety diagnosis	692	11.9	727	7.1	1,419	8.9
Diabetes mellitus	284	4.9	545	5.4	829	5.2
Hypertension	817	14.0	1,792	17.6	2,609	16.3
Statin prescription	340	5.8	806	7.9	1,146	7.2
Elevated total cholesterol	224	3.8	508	5.0	732	4.6
Anti-depressant use	709	12.2	781	7.7	1,490	9.3
Baseline CD4 T-cell count, cells/μL[b]	349 (173.0–524.0)		323.0 (178.0–461.0)		332.0 (177.0–485.0)	
Baseline viral load (log10), copies/mL[b]	4.6 (3.9–5.1)		4.6 (3.9–5.1)		4.6 (3.9–5.1)	
Calendar year at initiation[b]	2014 (2013–2015)		2011 (2010–2012)		2012 (2010–2014)	

Abbreviations: AIDS, acquired immune deficiency syndrome; EFV, efavirenz; HIV, human immunodeficiency virus; InSTI, integrase strand transfer inhibitor.

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

[b] Values are expressed as median (interquartile range).

[c] Weight (kg)/height $(m)^2$.

dolutegravir, and 3,361 (57.7%) elvitegravir-cobicistat. The composite outcome was experienced by 1,052 (10.3%) patients in the efavirenz-initiation group, where 650 (61.8%) had an AIDS-defining illness, 192 (18.3%) had a serious non-AIDS event, and 210 (20.0%) died from any cause. In the InSTI-initiation group, the composite outcome was experienced by 433 (7.4%) patients, where 281 (64.9%) had an AIDS-defining illness, 64 (14.8%) had a serious non-AIDS event, and 88 (20.3%) died from any cause. Additionally, 2,120 (20.8%) patients who initiated an efavirenz-based regimen and 752 (12.9%) patients who initiated an InSTI-based regimen were lost to follow up. There were 35,849 patient-years in the efavirenz group and 12,893 patient-years in the InSTI group.

The 5-year efavirenz risk estimate of the composite outcome was 12.3% (95% CI: 11.5, 13.1), and the InSTI risk estimate was 12.6% (95% CI: 11.0, 14.2). Under the observed treatment rule, the risk estimate was 12.5% (11.6, 13.4). The 5-year cross-validated risk estimates for the estimated optimal treatment rules found by each candidate precision medicine method is shown in Table 2. The absolute risk reduction of each treatment rule compared with recommending all patients an efavirenz-based regimen and recommending all patients an InSTI-based regimen is also shown in Table 2. Two of the 4 genetic algorithm methods produced the smallest risk estimates (or equivalently, the largest absolute risk differences) with similar standard errors. The smallest risk estimate was 10.0% (95% CI: 8.6, 11.3) using the SAGA method. The corresponding risk difference was 2.3% (95% CI: 0.9, 3.8) when comparing the estimated optimal treatment rule with efavirenz-only and 2.6% (95% CI: 1.0, 4.2) when comparing with InSTI-only. Compared with the observed treatment rule, the estimated optimal treatment rule reduced estimated risk by 2.5% (95% CI: 1.3, 3.7). These absolute risk reductions are equivalent to about an 18%–21% relative risk reduction.

Figure 1 shows the cross-validated risk curves over 5 years for the efavirenz-only, InSTI-only, observed treatment

Table 2. Summary of Implemented Statistical Methods for Precision Medicine and 5-Year Cross-Validated Risk[a] Estimates Using Data From the North American AIDS Cohort Collaboration on Research and Design, 2009–2016

Method	Details	Risk[a] at 5 years, %	Absolute Risk Difference vs. EFV[b]	Absolute Risk Difference vs. InSTI[b]	Absolute Risk Difference vs. Observed[b]
Outcome regression					
Penalized Cox regression	L2 penalty (ridge)	11.7	0.6	0.9	0.8
	L1 penalty (lasso)	12.1	0.3	0.6	0.4
	Mixed penalty (elastic net)	12.5	−0.2	0.1	−0.1
Causal survival forests	Number of trees was 5,000; 7 covariates considered at each split	12.2	0.1	0.4	0.3
Direct value search					
Genetic algorithm	Nonsmooth, nonaugmented	11.8	0.5	0.8	0.7
	Smooth, nonaugmented	10.5	1.8	2.1	2.0
	Nonsmooth, augmented	11.4	1.0	1.2	1.1
	Smooth, augmented	10.0	2.3	2.6	2.5
Outcome-weighted learning	Tuning parameter chosen by cross-validation	12.0	0.3	0.6	0.5
Residual-weighted learning	Tuning parameter chosen by cross-validation, accelerated failure time model chosen for outcome model	12.0	0.3	0.6	0.5
Indiscriminate					
All EFV[b]	N/A	12.3	0.0	0.3	0.2
All InSTI[b]	N/A	12.6	−0.3	0.0	−0.1
Observed[b]	N/A	12.5	−0.2	0.1	0.0

Abbreviations: AIDS, acquired immune deficiency syndrome; EFV, efavirenz; InSTI, integrase strand transfer inhibitor.
[a] Risk is with respect to risk of the composite of AIDS-defining illnesses, serious non-AIDS events, and all-cause mortality.
[b] EFV and InSTI refer to the indiscriminate treatment rule where all patients are recommended an EFV-based drug regimen or all patients are recommended an InSTI-based drug regimen, where both regimens include the same backbone of tenofovir disoproxil fumarate (or tenofovir alafenamide) and emtricitabine. Observed refers to the treatment rule where patients are recommended their observed drug regimen.

rule, and SAGA-estimated optimal treatment rule. The optimal treatment risk curve tracks closely with the efavirenz risk curve initially but then gradually separates after about 2 years. Web Figure 1 shows the risk curves separately with 95% CIs.

Estimated cumulative incidence functions for each of the 3 components of the composite outcome are shown in Web Figure 2. The estimated optimal treatment rule had similar or lower estimated 5-year risks for each of the composition outcome components.

Since the SAGA method produced the smallest cross-validated risk estimate, it was applied to the entire data set to estimate the optimal treatment rule, which is reported in Table 3 (see Web Tables 5–12 for the optimal treatment rules estimated by the other methods, except causal survival forests, which do not have an explicit form for a treatment rule; Web Table 13 shows summary statistics of the optimal treatment rule estimated by causal survival forests). The estimated optimal treatment rule is a linear function of the covariates. Continuous covariates were centered and standardized by 2 standard deviations (without stratification by treatment group) to allow for interpretation on roughly the same scale as the binary covariates (11). The treatment rule is coded so that positive values correspond to recommending an InSTI-based regimen, and an efavirenz-based regimen is recommended otherwise. The SAGA-estimated optimal treatment rule recommended an InSTI-based regimen for about 57% of patients and differed from the observed treatments for about 54% of patients.

To facilitate use of this study, an Excel (Microsoft) spreadsheet that encodes the estimated optimal treatment rule is provided in Web Appendix 2. As an example, consider a hypothetical male patient with hepatitis C infection but no other health conditions or behaviors listed in the covariate column of Table 3, and whose age, body mass index, viral load, and CD4 T-cell count equal the average of the NA-ACCORD cohort. For this patient, the treatment rule would sum the hepatitis C coefficient and the intercept coefficient, $-0.451 + 0.153 = -0.298$, which is negative, so an efavirenz-based regimen would be recommended.

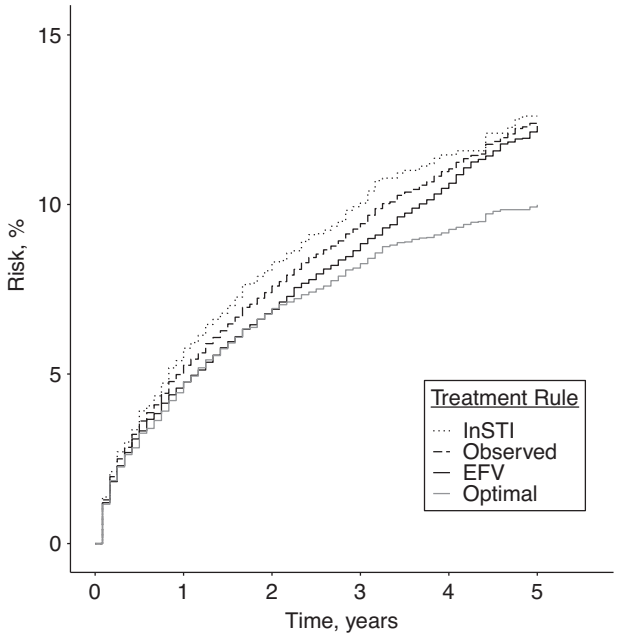

Figure 1. Estimated 5-year cross-validated risk curves when recommending an efavirenz (EFV)-based regimen to all patients, an integrase strand transfer inhibitor (InSTI)-based regimen to all patients, according to the observed treatments, and according to the optimal treatment rule estimated by the genetic algorithm method with smoothing and augmentation (see Web Appendix 3 for details), using data from the North American AIDS Cohort Collaboration on Research and Design, 2009–2016. Risk is with respect to the composite of acquired immune deficiency syndrome (AIDS)-defining illnesses, serious non-AIDS events, and all-cause mortality. The optimal treatment rule is a function that takes in patient characteristics and outputs a recommendation for an EFV- or InSTI-based regimen to minimize risk of the composite outcome.

Table 3. Estimated Coefficients of the Optimal Treatment Rule[a] That Minimized the 5-Year Composite Disease Risk of AIDS-Defining Illnesses, Serious Non-AIDS Events, and All-Cause Mortality Using Data From the North American AIDS Cohort Collaboration on Research and Design, 2009–2016

Covariate[b]	Coefficient[c]
Hepatitis C infection	−0.451
Statin prescription	−0.381
Diabetes mellitus	−0.286
Men who have sex with men	−0.246
Risky heterosexual sex	−0.213
CD4 T-cell count	−0.198
Injection drug use	−0.180
Female sex	−0.126
Elevated total cholesterol	−0.071
Hepatitis B infection	0.017
Body mass index[d]	0.032
Age	0.091
Viral load	0.102
Hispanic/Latino	0.124
Black race	0.126
Anxiety	0.126
Intercept	0.153
Depression	0.209
Hypertension	0.259
Antidepressant use	0.289
Prior AIDS diagnosis	0.293

Abbreviation: AIDS, acquired immune deficiency syndrome.

[a] The treatment rule is linear in the covariates and was estimated by the genetic algorithm method with smoothing and augmentation. Negative values weight the treatment rule towards recommending EFV-based regimens while positive values weight the treatment rule towards recommending InSTI-based regimens. For a given patient, the estimated treatment rule recommends an InSTI-based regimen if the sum of all covariate-coefficient products is greater than or equal to 0, and the treatment rule recommends an EFV-based regimen otherwise.

[b] Age, body mass index, CD4 T-cell count, and viral load (log base 10 transformed) are continuous, and all other covariates are binary. Continuous covariates were centered and scaled by 2 standard deviations to allow for comparisons with the binary covariates (11).

[c] Coefficients were rounded to 3 decimal places.

[d] Weight (kg)/height (m)^2.

DISCUSSION

Applying statistical methods for precision medicine to observational data from a large collaboration of HIV cohorts in the United States and Canada suggests that tailoring efavirenz and InSTI regimens based on patient characteristics could reduce absolute 5-year disease risk by about 2%, compared with recommending the same regimen for all patients. Further studies that consider related clinical outcomes or different populations are needed to validate these results. For instance, complementary evidence would be provided by a randomized clinical trial with biomarker outcomes, such as viral load and CD4 count, where one arm of the trial follows standard of care and the other arm follows the estimated optimal treatment rule.

The strengths of this study include the large, representative data set and the use of modern statistical tools grounded in a rigorous causal framework. The NA-ACCORD data followed a diverse group of patients over a prolonged study duration to observe outcomes of direct interest. Various precision medicine methods to estimate the optimal treatment rule were considered, and a sample splitting procedure was used to mitigate the effect of overfitting. The methods implemented utilize consistent estimators of the causal effects of different treatment rules, provided the assumptions discussed below hold.

Assumptions of this analysis included (but are not limited to) no unmeasured confounding and treatment positivity. With observational data, the no-unmeasured-confounding assumption is generally untestable and will be violated if there is an unmeasured covariate that, conditional on observed covariates, influences both treatment assignment and the outcome. For instance, HIV genotype (not

considered in this analysis) may have affected the treatment assignment mechanism and interacted with treatment to influence the composite outcome. Future analyses may avoid the no unmeasured-confounding-assumption by using data from clinical trials or methods that exchange the no-unmeasured-confounding assumption for other assumptions that may be more plausible under certain data-generating mechanisms. Treatment positivity assumes that patients within each covariate stratum have positive probability of receiving either treatment. Comparison of the marginal distributions of each baseline covariate according to treatment group status did not suggest violation of this assumption. See Web Appendix 3 and Web Figure 3 for further discussion of the assumptions employed in this analysis.

This study has limitations. Although the magnitude of the estimated risk reduction from tailoring treatments is potentially clinically relevant, the corresponding CIs were wide, suggesting moderate uncertainty about the estimate. There may also be residual model overfitting; although a cross-validation procedure was used to reduce the amount of overfitting, it does not eliminate the possibility. In other words, the reported estimates may be too optimistic. Additionally, since the data are observational, neither the treatment assignment or censoring mechanism are known, and it is possible that there was unmeasured confounding, model misspecification, or measurement error. This analysis also did not consider the possibility that there are substantial differences between drug regimens within the InSTI class. This study focused only on the effect of the initial treatment regimen; future analyses could investigate optimal treatment rules that may vary treatment recommendations at different time points, conditional on patient history up to that point.

Current guidelines for HIV treatment recommend InSTI-based regimens for most patients who are ART-naive (1–3, 6). Subject to additional evidence from similar analyses on independent data sets or (ideally) a confirmatory clinical trial, the results presented in this paper offer evidence for potential updating of these guidelines. The estimated optimal treatment rule from this study uses commonly collected covariate data and would be straightforward for clinicians to implement. We note that this tool should not replace expert clinical judgment since this study does not consider some relevant outcomes, such as side effect profiles, costs of treatments, or drug resistance. Rather, the estimated treatment rule reported here should be used in conjunction with consideration of those other outcomes. For instance, if concerns about side effects or costs were deemed to be roughly equal for both treatments, the reported estimated treatment rule could be used to make recommendations. The estimated optimal treatment rule would serve as an additional tool for medical practitioners and could help standardize care.

ACKNOWLEDGMENTS

Author affiliations: Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States (Michael Jetsupphasuk, Michael G. Hudgens); Department of Epidemiology of Microbial Diseases, Yale School of Public Health, Yale University, New Haven, Connecticut, United States (Haidong Lu); Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States (Stephen R. Cole, Jessie K. Edwards, Adaora A. Adimora, Joseph J. Eron); Division of Infectious Diseases, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States (Adaora A. Adimora, Joseph J. Eron); Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, United States (Keri N. Althoff, Raynell Lang); Division of Research, Kaiser Permanente Northern California, Oakland, California, United States (Michael J. Silverberg); Vanderbilt University Medical Center, Nashville, Tennessee, United States (Peter F. Rebeiro, Timothy R. Sterling); Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada (Viviane D. Lima); Division of Infectious Diseases, Emory University School of Medicine, Atlanta, Georgia, United States (Vincent C. Marconi); Department of Global Health, Emory University Rollins School of Public Health, Georgia, United States (Vincent C. Marconi); Mid-Atlantic Permanente Research Institute, Kaiser Permanente Mid-Atlantic States, Rockville, Maryland, United States (Michael A. Horberg); Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada (M. John Gill, Raynell Lang); Department of Medicine, School of Medicine, University of Washington, Seattle, Washington, United States (Mari M. Kitahata); Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States (Richard D. Moore, Kelly Gebo); and Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, United States (Charles Rabkin).

This work was supported by the National Institutes of Health (grants U01AI069918, F31AI124794, F31DA037788, G12MD007583, K01AI093197, K01AI131895, K23EY013707, K24AI065298, K24AI118591, K24DA000432, KL2TR000421, N01CP01004, N02CP055504, N02CP91027, P30AI027757, P30AI027763, P30AI027767, P30AI036219, P30AI050409, P30AI050410, P30AI094189, P30AI110527, P30MH62246, R01AA016893, R01DA011602, R01DA012568, R01 AG053100, R24AI067039, U01AA013566, U01AA020790, U01AI038855, U01AI038858, U01AI068634, U01AI068636, U01AI069432, U01AI069434, U01DA03629, U01DA036935, U10EY008057, U10EY008052, U10EY008067, U01HL146192, U01HL146193, U01HL146194, U01HL146201, U01HL146202, U01HL146203, U01HL146204, U01HL146205, U01HL146208, U01HL146240, U01HL146241, U01HL146242, U01HL146245, U01HL146333, U24AA020794,U54MD007587, UL1RR024131,

UL1TR000004, UL1TR000083, Z01CP010214, and Z01CP010176; grants P30AI050410 to M.G.H., S.R.C., J.K.E., A.A.A., and J.J.E.; R01AI157758 to M.J., M.G.H., S.R.C., J.K.E., A.A.A., and J.J.E.; and P30AI050409 to V.C.M.) and contracts CDC-200-2006-18797 and CDC-200-2015-63931 from the US Centers for Disease Control and Prevention; contract 90047713 from the US Agency for Healthcare Research and Quality; contract 90051652 from the US Health Resources and Services Administration; grants CBR-86906, CBR-94036, HCP-97105 and TGF-96118 from the Canadian Institutes of Health Research; Ontario Ministry of Health and Long Term Care; and the Government of Alberta, Canada. Additional support was provided by the National Institute of Allergy And Infectious Diseases (NIAID), National Cancer Institute (NCI), National Heart, Lung, and Blood Institute (NHLBI), *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), National Human Genome Research Institute (NHGRI), National Institute for Mental Health (NIMH) and National Institute on Drug Abuse (NIDA), National Institute on Aging (NIA), National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Neurological Disorders and Stroke (NINDS), National Institute of Nursing Research (NINR), National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institute on Deafness and Other Communication Disorders (NIDCD), and National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

Data are from the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) of the International Epidemiology Databases to Evaluate AIDS (IeDEA). A concept sheet must be submitted, reviewed, and approved by the Executive and Steering Committees of NA-ACCORD before data access/data sharing are allowed. This is a requirement of all participating cohort institutional review board (IRB) approvals, as well as the study umbrella IRB approval from the Johns Hopkins University School of Medicine, ensuring secure, timely, and ethical sharing of cohort data. Due to ethical restrictions, data are available upon request. The NA-ACCORD's data practices have been outlined in the literature (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2820873/). Interested researchers may submit requests for data to the project manager, who will direct the collaborative process through the Administrative, Data Management, and Epidemiology/Biostatistics Cores (afreeman@jhu.edu). For researchers with access to the data, code to reproduce the main results of this study is available at https://github.com/michaeljets/pm_hiv.

The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the NIH.

Conflict of interest: M.J.G. has received honoraria as ad hoc member of HIV national advisory boards to Merck, Gilead, and ViiV Health. V.C.M. has received investigator-initiated research grants (to the institution) and consultation fees (both unrelated to the current work) from Eli Lilly, Bayer, Gilead Sciences, and ViiV. The other authors report no conflicts.

REFERENCES

1. Saag MS, Benson CA, Gandhi RT, et al. Antiretroviral drugs for treatment and prevention of HIV infection in adults: 2018 recommendations of the International Antiviral Society–USA Panel. *JAMA.* 2018;320(4):379–396.
2. World Health Organization. Update of recommendations on first- and second-line antiretroviral regimens. https://www.who.int/publications-detail-redirect/WHO-CDS-HIV-19.15. Accessed June 7, 2021.
3. National Institutes of Health. What to start: initial combination regimens for the antiretroviral-naive patient. https://clinicalinfo.hiv.gov/en/guidelines/adult-and-adolescent-arv/what-start-initial-combination-regimens-antiretroviral-naive. Accessed June 7, 2021.
4. Lu H, Cole SR, Westreich D, et al. Clinical effectiveness of integrase strand transfer inhibitor–based antiretroviral regimens among adults with human immunodeficiency virus: a collaboration of cohort studies in the United States and Canada. *Clin Infect Dis.* 2021;73(7):e1408–e1414.
5. Cole SR, Edwards JK, Hall HI, et al. Incident AIDS or death after initiation of human immunodeficiency virus treatment regimens including raltegravir or efavirenz among adults in the United States. *Clin Infect Dis.* 2017;64(11):1591–1596.
6. Horberg MA, Oakes AH, Hurley LB, et al. Association of raltegravir use with long-term health outcomes in HIV-infected patients: an observational post-licensure safety study in a large integrated healthcare system. *HIV Clin Trials.* 2018;19(5):177–187.
7. Kosorok MR, Laber EB. Precision medicine. *Annu Rev Stat Appl.* 2019;6:263–286.
8. Tsiatis AA, Davidian M, Holloway ST, et al. *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. Boca Raton, FL: Chapman and Hall/CRC; 2020.
9. INSIGHT START Study Group. Initiation of antiretroviral therapy in early asymptomatic HIV infection. *N Engl J Med.* 2015;373(9):795–807.
10. Babiker AG, Emery S, Fätkenheuer G, et al. Considerations in the rationale, design and methods of the Strategic Timing of AntiRetroviral Treatment (START) study. *Clin Trials.* 2013;10(1 suppl):S5–S36.
11. 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.
12. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics.* 2012;28(1):112–118.
13. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B Methodol.* 1996;58(1):267–288.
14. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Series B Stat Methodology.* 2005;67(2):301–320.
15. Cui Y, Kosorok MR, Wager S, et al. Estimating heterogeneous treatment effects with right-censored data via causal survival forests [preprint]. *arXiv*. 2020. https://doi.org/10.48550/arXiv.2001.09887 Accessed February 21, 2021.
16. Jiang R, Lu W, Song R, et al. On estimation of optimal treatment regimes for maximizing t-year survival probability. *J R Stat Soc Series B Stat Methodology.* 2017;79(4):1165–1185.
17. Zhao YQ, Zeng D, Laber EB, et al. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika.* 2015;102(1):151–168.

18. Zhou X, Mayer-Hamblett N, Khan U, et al. Residual weighted learning for estimating individualized treatment rules. *J Am Stat Assoc.* 2017;112(517):169–187.
19. Zhou J, Zhang J, Lu W, et al. On restricted optimal treatment regime estimation for competing risks data. *Biostatistics.* 2021;22(2):217–232.
20. Jiang X, Nelson AE, Cleveland RJ, et al. Precision medicine approach to develop and internally validate optimal exercise and weight-loss treatments for overweight and obese adults with knee osteoarthritis: data from a single-center randomized trial. *Arthritis Care Res.* 2021;73(5): 693–701.