A Comparison of Seven Cox Regression-Based Models to Account for Heterogeneity Across Multiple HIV Treatment Cohorts in Latin America and the Caribbean

Mark J. Giganti,¹ Paula M. Luz,² Yanink Caro-Vega,³ Carina Cesar,⁴ Denis Padgett,⁵ Serena Koenig,^{6,7} Juan Echevarria,⁸ Catherine C. McGowan,⁹ and Bryan E. Shepherd¹

Abstract

Many studies of HIV/AIDS aggregate data from multiple cohorts to improve power and generalizability. There are several analysis approaches to account for cross-cohort heterogeneity; we assessed how different approaches can impact results from an HIV/AIDS study investigating predictors of mortality. Using data from 13,658 HIV-infected patients starting antiretroviral therapy from seven Latin American and Caribbean cohorts, we illustrate the assumptions of seven readily implementable approaches to account for across cohort heterogeneity with Cox proportional hazards models, and we compare hazard ratio estimates across approaches. As a sensitivity analysis, we modify cohort membership to generate specific heterogeneity conditions. Hazard ratio estimates varied slightly between the seven analysis approaches, but differences were not clinically meaningful. Adjusted hazard ratio estimates for the association between AIDS at treatment initiation and death varied from 2.00 to 2.20 across approaches that accounted for heterogeneity; the adjusted hazard ratio was estimated as 1.73 in analyses that ignored across cohort heterogeneity. In sensitivity analyses with more extreme heterogeneity, we noted a slightly greater distinction between approaches. Despite substantial heterogeneity between cohorts, the impact of the specific approach to account for heterogeneity was minimal in our case study. Our results suggest that it is important to account for across cohort heterogeneity in analyses, but that the specific technique for addressing heterogeneity may be less important. Because of their flexibility in accounting for cohort heterogeneity, we prefer stratification or meta-analysis methods, but we encourage investigators to consider their specific study conditions and objectives.

Introduction

MANY OBSERVATIONAL HIV/AIDS STUDIES are performed with data pooled from multiple cohorts. Examples include the International Epidemiologic Databases to Evaluate AIDS (IeDEA),^{1,2} EuroCoord,³ and the Antiretroviral Therapy Cohort Collaboration (ART-CC),⁴ to name just a few. The benefits of multi-cohort studies are numerous, including the ability to study rare events and the improved generalizability of results.

Cohort membership can be highly associated with patient-level data and outcomes in manners that are often not adequately captured by patient-level covariates. For example, a multicohort study of HIV-infected patients starting combination antiretroviral therapy (ART) in Europe and North America recently reported considerable differences in rates of mortality and AIDS across cohorts, even after adjusting for factors commonly considered predictive of death and disease progression.⁵ For analyses of time-to-event outcomes, there are many strategies to account for cross-cohort heterogeneity,

¹Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee.

²Instituto de Pesquisa Clínica Evandro Chagas, Fiocruz, Rio de Janeiro, Brazil.

³Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubiran, Mexico City, Mexico.

⁴Investigaciones Clínicas, Fundación Huésped, Buenos Aires, Argentina.

⁵Instituto Hondureño de Seguridad Social and Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras.

⁶Les Centres GHESKIO, Port-au-Prince, Haiti.

⁷Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

⁸Instituto de Medicina Tropical Alexander von Humboldt of the Universidad Peruana Cayetano Heredia and Hospital Nacional Cayetano Heredia, Lima, Perú.

⁹Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee.

including directly adjusting for study cohort in the model, fitting frailty or random effects models,⁶ stratifying by cohort,⁷ accounting for correlation using robust variance estimators, or fitting separate models in each cohort and aggregating summary measures.⁸

Several studies in the statistical literature have compared the performance of specific approaches under various conditions.^{9–12} However, to our knowledge, there are no published studies that investigate various analytic approaches to adjust for cross-cohort heterogeneity among pooled HIV cohorts. In the HIV literature, there is no consensus on which approach to use; recent studies have used different approaches, including fixed effect terms,¹³ random effect terms,¹⁴ stratification,^{15,16} meta-analyses,¹⁷ and no adjustment.^{18,19} In our own studies, we have debated which modeling strategy to use and have wondered how results might vary across methods.

Using data from a multicohort observational study of HIV-infected patients in Latin America and the Caribbean, we investigate the impact of different strategies for modeling cross-cohort heterogeneity on study estimates, specifically estimates of the association between AIDS at ART initiation and mortality. The HIV epidemic, level of economic development, and background mortality rates vary extensively across the region^{20,21}; there is also heterogeneity in the classification of AIDS. We apply seven readily implementable Cox regression methods, illustrate their implicit assumptions, and compute and compare resulting hazard ratios estimates. As a sensitivity analysis, we permute cohort membership in a subsample of patients to explore how estimates can be affected by different cohort compositions.

Materials and Methods

Study cohort

The Caribbean, Central, and South American Network for HIV Epidemiology (CCASAnet) is a collaboration of clinics from seven countries and is Region 2 of IeDEA. CCASAnet cohorts contributing data to this study were Centro Médico Huésped, Buenos Aires, Argentina (CMH-Argentina); Instituto de Pesquisa Clínica Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil (IPEC-Brazil); Fundación Arriarán, Santiago, Chile (FA-Chile); Le Groupe Haitien d'Etude du Sarcome de Kaposi et des Infections Opportunistes, Port-au-Prince, Haiti (GHESKIO-Haiti); Instituto Hondureño de Seguridad Social and Hospital Escuela, Tegucigalpa, Honduras (IHSS/HE-Honduras); Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico (INCMNSZ-Mexico); and Instituto de Medicina Tropical Alexander von Humboldt, Lima, Perú (IMTA vH-Peru). Data were collected at each site and sent to the CCASAnet Data Coordinating Center at Vanderbilt University (Nashville, Tennessee) for harmonization. Institutional review board approval was obtained from each site and Vanderbilt University.

Antiretroviral-naïve adults (>17 years) initiating their first ART regimen on or after January 1, 2000 were included in this study. The database closing date was April 30, 2011 for FA-Chile and April 30, 2013 for all remaining cohorts. The primary outcome was all-cause mortality. At all sites, patient deaths were ascertained using clinical records (if death occurred at a hospital) or notification by relatives. In IPEC-Brazil, FA-Chile, INCMNSZ-Mexico, and IMTA vH-

Peru, national death registry databases were also checked annually. A patient was classified as lost to follow-up (LTFU) if they had no clinical visit within the year preceding the database closing date.

The primary predictor was stage of disease at ART initiation, categorized as AIDS or not AIDS. The classification of AIDS was heterogeneous across cohorts. AIDS was defined as WHO stage IV at GHESKIO-Haiti and IMTA vH-Peru; CDC stage C at IHSS/HE-Honduras and INCMNSZ-Mexico; and a designation of AIDS (yes/no) at CMH-Argentina, IPEC-Brazil, and FA-Chile (which did not distinguish between clinical events and CD4 <200). Nadir CD4⁺ cell count (CD4) at ART initiation was defined as the lowest CD4 measurement prior to or no more than 7 days after ART initiation.

Statistical analyses

The association between AIDS at ART initiation and mortality was modeled using a Cox model, with the hazard of death at time t , $\lambda(t)$, being equal to $\lambda_0(t)\exp(X_1\beta_1 + \dots + X_P\beta_P)$, where $\lambda_0(t)$ is the baseline hazard (hazard at t when all covariates are at reference levels), $\{X_1, \dots, X_P\}$ represent the indicator of AIDS (yes/no) and $P - 1$ other covariates, and $\{\beta_1, \dots, \beta_P\}$ are the natural logarithm of the covariate hazard ratios. The parameters $\{\beta_1, \dots, \beta_P\}$ are sometimes referred to as “treatment effects” (although X need not be a treatment and the parameters may represent associations, not causal effects).

Within the Cox model, we considered seven possible strategies to account for cross-cohort heterogeneity: (1) ignore study cohort information, which implicitly assumes that the baseline hazard function and hazard ratios (HRs) are the same for each cohort (*naïve approach*); (2) account for correlation between patients within cohorts in computation of standard errors using robust variance estimation, but otherwise assume baseline hazard function and HRs are the same for each cohort (*robust marginal approach*); (3) directly adjust for cohort as a covariate and estimate corresponding hazard ratios for each cohort (*fixed effects approach*); (4) adjust for cohort in the model as a random effect assumed to come from a normal distribution (*random effects approach*); (5) adjust for cohort as a random effect and allow the association between AIDS at ART initiation and mortality to vary by cohort (*random cohort/treatment effects approach*); (6) allow the baseline hazards to vary by cohort (*stratified approach*); and (7) fit separate models for each cohort and then aggregate the HR estimates using standard random effects meta-analysis techniques that use inverse variance weighting to aggregate estimates across cohorts (*meta-analysis approach*).²²

A more detailed description of each modeling technique is given in the Supplementary Material (Supplementary Fig. S1 and Supplementary Tables S1, S2, S3, and S4; Supplementary Data are available online at www.liebertpub.com/aid). The random effects approach was also fit assuming random effects came from a gamma distribution; results were similar to those assuming normality and are not repeated here. We assumed proportional hazards for AIDS and all other covariates in all models unless otherwise specified. The random effects, fixed effects, and random cohort/treatment effects approaches implicitly assume the hazards are proportional between cohorts whereas the stratified and meta-analysis approaches relax this assumption.

We first fit separate Cox models using each of the seven approaches, including AIDS at ART initiation but adjusting for no other covariates. Patients with no documented death were censored at their last clinical visit date. We estimated the hazard of death over time by cohort for those with and without AIDS at ART initiation and plotted a loess smoothed curve²³ of these estimated hazards of death. Hazard ratio estimates for AIDS at ART initiation versus no AIDS (reference category) for each modeling strategy, and 95% confidence intervals, were computed.

A second set of analyses applied all approaches while adjusting for patient age at ART initiation, sex, nadir CD4 at ART initiation, class of initial ART regimen (nucleoside reverse transcriptase inhibitor based, protease inhibitor based, or other), and the year of ART initiation. In these adjusted analyses, missing data (14% of AIDS diagnoses and 7% of CD4 measurements) were assumed to be missing at random given the other variables and were imputed using multiple imputation.²⁴ We report HRs and corresponding 95% confidence intervals for each approach. To assess relative goodness of fit, we also compared the Akaike Information Criterion (AIC) of each approach.²⁵ The proportional hazards assumption was tested by computing scaled Schoenfeld residuals separately for each predictor and then performing a test of correlation with time.²⁶ To explore time-dependent associations between AIDS at ART initiation and mortality, we fit separate adjusted Cox models that limit follow-up to 1-year intervals.

Sensitivity analyses were performed to investigate how results would change if there were different levels of cross-cohort heterogeneity, specifically: (1) no heterogeneity between cohorts, (2) more heterogeneity between cohorts' baseline hazards, (3) more heterogeneity between cohort-specific HRs for AIDS at ART initiation, and (4) more cohorts. Different levels of cross-cohort heterogeneity were generated by strategically reassigning a single variable, cohort, for certain patients. Baseline characteristics and outcome events remained identical for every patient in the original dataset. The overall sample size, as well as the relative sample size of each cohort (except for scenario 4), was unchanged. For each modified dataset, we performed multivariable analyses using each analysis strategy and report HRs and confidence intervals for AIDS at ART initiation.

To generate scenario (1), we reassigned cohorts for each patient by randomly permuting the original cohort assignments. To generate (2), we reassigned 1,820 patient records from four cohorts according to their length of follow-up and whether a death was observed. The objectives of these changes were to have at least one cohort with an increasing hazard over time, to have at least one cohort with a flat hazard, and to steepen the slope of the decreasing hazards for the other modified cohorts. To generate (3), we randomly selected 250 patient records from one cohort with AIDS who had an observed death and switched them to the second cohort; similarly, we randomly selected 250 patients from the second cohort with AIDS who did not have a recorded death and switched them to the first. Finally, to generate (4), we randomly divided original cohorts into subcohorts of approximately constant size. We generated four modified datasets for this scenario; the total number of cohorts ranged from 23 to 271, corresponding to cohort sizes of approximately 500 to 50 patients, respectively. A complete list of

modifications for each sensitivity analysis is included in the supplementary materials.

All analyses were performed using R Version 2.15. The analysis code is posted at <http://biostat.mc.vanderbilt.edu/ArchivedAnalyses>.

Results

Data from 13,658 HIV-infected ART initiating patients were included in this study, with variable patient numbers across cohorts: 823 from CMH-Argentina, 1,765 from IPEC-Brazil, 1,012 from FA-Chile, 6,300 from GHESKIO-Haiti, 900 from IHSS/HE-Honduras, 875 from INCMNSZ-Mexico, and 1,983 from IMTAvH-Peru. In the combined cohort, 32% of patients had AIDS at ART initiation, but this also varied across cohorts: 7% in CMH-Argentina, 24% in GHESKIO-Haiti, 27% in IMTAvH-Peru, 31% in IHSS/HE-Honduras, 52% in IPEC-Brazil, 55% in INCMNSZ-Mexico, and 57% in FA-Chile. The median follow-up was 3.7 years (interquartile range 1.9–6.0 years), varying from 3.1 years at IMTAvH-Peru to 5.2 years at FA-Chile. The proportions of patients LTFU also varied by cohort from 59% at CMH-Argentina to 5% at INCMNSZ-Mexico. Overall, 1,319 (9.7%) died: 13 from CMH-Argentina, 161 from IPEC-Brazil, 102 from FA-Chile, 756 from GHESKIO-Haiti, 110 from IHSS/HE-Honduras, 46 from INCMNSZ-Mexico, and 131 from IMTAvH-Peru. More details are found elsewhere.^{21,27} A more complete summary of patient characteristics by cohort is in Supplementary Table S1.

Figure 1 shows the estimated hazards of death over time by study cohort, and demonstrates the substantial across cohort heterogeneity. The estimated hazard is decreasing over the entire study period for some cohorts (e.g., GHESKIO-Haiti, IHSS/HE-Honduras, IMTAvH-Peru) whereas the hazard appears to plateau and either stay the same or increase for other cohorts (e.g., IPEC-Brazil, FA-Chile, INCMNSZ-Mexico).

FIG. 1. Unadjusted estimated hazard of death following HIV treatment initiation for seven cohorts from the Caribbean and the Central and South American Network for HIV Epidemiology ($N = 13,658$).

We also note the unique hazard for CMH-Argentina, where the hazard is lower except for an unusual spike around 1 year likely due to small numbers of events.

Figure 2 shows the model-based predicted hazards of death by cohort and by AIDS status at ART initiation for each of the seven strategies to account for cross-cohort heterogeneity. The naive approach implicitly assumes the same hazard per cohort (Fig. 2A); the estimated hazard function for the robust marginal approach is identical. Under the fixed effects approach (Fig. 2B), the estimated hazards now differ by cohort; the curves are parallel (i.e., the hazards are assumed to be proportional) across cohorts and the hazard ratio comparing those with and without AIDS is the same for each. The plot corresponding to post hoc hazard estimates from the random effects approach (Fig. 2C) has a very similar, although not identical, pattern to that of the fixed effects approach. Figure 2D shows estimated hazards using the random cohort/treatment effects approach. Curves are parallel, although unlike the previous approaches, the hazard ratios for AIDS at ART initiation are allowed to vary by cohort. Figure 2E shows

estimates from the stratified approach. Similar to the marginal, fixed cohort effects, and random cohort effects approaches, the relative distance between those with and without AIDS at ART initiation is assumed to be the same for each cohort. However, hazards are no longer parallel across cohorts because each has its own baseline hazard function. Finally, Fig. 2F shows the estimated hazard functions when separate models are fit to each cohort, thereby allowing the baseline hazard and the association with AIDS to vary by cohort.

Table 1 reports HRs and corresponding 95% confidence intervals for death by AIDS at ART initiation using the seven different approaches to account for heterogeneity after adjusting for age, sex, nadir CD4, year of initiation, and initial regimen. For the naive and robust marginal approaches, hazard ratio estimates were identical. Accounting for the correlation within cohorts with the robust marginal approach widened confidence intervals, although not substantially. Results from the other five approaches tended to differ from the marginal approaches. However, among these five approaches, we

FIG. 2. (A–F) Estimated hazard of death following HIV treatment initiation by cohort and AIDS status based on various modeling strategies. The *dark line* corresponds to the hazard of death among those without AIDS and the *gray line* corresponds to the hazard of death among those with AIDS ($N=13,658$).

TABLE 1. HAZARD RATIO ESTIMATES AND 95% CONFIDENCE INTERVALS FOR DEATH USING THE DIFFERENT STRATEGIES FOR ACCOUNTING FOR HETEROGENEITY BETWEEN SEVEN COHORTS FROM THE CARIBBEAN, CENTRAL, AND SOUTH AMERICAN NETWORK FOR HIV EPIDEMIOLOGY ($N=13,658$)

	<i>AIDS at treatment initiation</i>
Naive marginal	1.73 (1.53–1.96)
Robust marginal	1.73 (1.37–2.19)
Fixed effects	2.00 (1.75–2.29)
Random effects ^a	2.01 (1.76–2.29)
Random cohort/treatment effects ^a	2.13 (1.72–2.65)
Stratified	2.01 (1.76–2.30)
Meta-analysis	2.20 (1.64–2.95)

^aModel assumes a Gaussian random effects distribution.

All analyses were adjusted for age, sex, nadir CD4 count, year of initiation, initial ART regimen, and AIDS defining condition at ART initiation.

obtained only slightly different estimates, none of which would be considered clinically meaningful. The fixed and random effects analyses were nearly identical. Estimates that assumed a separate baseline hazard per cohort (stratified approach) were also very similar to those from the fixed and random effects analyses. Finally, allowing the AIDS effect to vary by cohort only slightly raised estimates and widened confidence intervals (random cohort/treatment effects and meta-analysis approaches).

The meta-analysis and stratified approaches had lower AIC (19,915 and 19,920, respectively) than the random effects approach (23,534), random cohort/treatment effects approach (23,535), fixed effects approach (23,537), and naive (23,743) and robust marginal approaches (23,743). There was also statistical evidence that the hazards were not proportional across cohorts ($p < 0.01$, fixed effects model). These results suggest that the meta-analysis and stratified approaches better fit the data.

Sensitivity analyses

In the analyses presented above, results were fairly similar regardless of the statistical modeling strategy. While inter-

esting for our dataset, we wondered what scenarios might lead to clinically different findings. By permuting cohort membership, we obtained hazard ratio estimates corresponding to settings with (1) no heterogeneity between cohorts, (2) more heterogeneity between cohorts' baseline hazards, (3) more heterogeneity between cohort-specific HRs for AIDS at ART initiation, and (4) more cohorts. Results from settings (1)–(3) are in Table 2; results from setting (4) are given in Table 3.

If there was no heterogeneity between cohorts, setting (1), the association between AIDS and the hazard of death was nearly identical for each analysis strategy, as expected (Table 2, column 1). The corresponding confidence intervals were also similar for every approach except for the robust marginal approach, suggesting that losses in efficiency by using more flexible models may be minimal.

Despite increasing the heterogeneity of the baseline hazards (Fig. 3A) when data were permuted to produce extreme heterogeneity in the baseline hazards under setting (2), we did not observe any clinically meaningful difference in hazard ratio estimates between the stratified and fixed/random effects approaches (Table 2, column 2). Figure 3B shows the estimated hazards under setting (3) with data permuted to produce extreme variation across cohorts in the HR estimates. Results using this modified dataset (Table 2, column 3) suggest a greater distinction in the estimates between the approaches that account for heterogeneity in hazard ratios (random cohort/treatment effects and meta-analysis approaches) and those that do not.

Table 3 shows estimates from the fixed effects, random effects, and stratified approaches when the number of cohorts increased. Estimates from the fixed effects and stratified approaches were similar regardless of the number of cohorts, but both differed from those of the random effects approach as the number of cohorts increased. The confidence intervals were slightly wider for the fixed effects approaches compared to the random effects approaches as the number of cohorts increased.

Discussion

In this study, we applied seven common strategies to account for heterogeneity between cohorts in analyses investigating the association between AIDS at ART initiation and

TABLE 2. HAZARD RATIO ESTIMATES AND 95% CONFIDENCE INTERVALS FOR DEATH USING THE DIFFERENT STRATEGIES FOR ACCOUNTING FOR HETEROGENEITY BETWEEN SEVEN COHORTS USING MODIFIED DATASET ($N=13,658$)

	<i>AIDS at treatment initiation</i>		
	<i>Random site</i>	<i>Extreme heterogeneity</i>	<i>Extreme treatment effect</i>
Naive marginal	1.73 (1.53–1.96)	1.73 (1.53–1.96)	1.73 (1.53–1.96)
Robust marginal	1.73 (1.58–1.90)	1.73 (1.31–2.29)	1.73 (0.54–5.53)
Fixed effects	1.73 (1.53–1.96)	1.71 (1.50–1.94)	1.35 (1.18–1.54)
Random effects ^a	1.73 (1.53–1.96)	1.71 (1.50–1.95)	1.36 (1.19–1.55)
Random cohort/treatment effects ^a	1.73 (1.52–1.97)	1.68 (1.22–2.32)	1.93 (0.92–4.07)
Stratified	1.73 (1.53–1.96)	1.65 (1.45–1.88)	1.36 (1.19–1.55)
Meta-analysis	1.72 (1.50–1.97)	1.65 (1.06–2.55)	2.10 (0.78–5.65)

^aModel assumes a Gaussian random effects distribution.

All analyses were adjusted for age, sex, nadir CD4 count, year of initiation, initial ART regimen, and AIDS defining condition at ART initiation.

TABLE 3. HAZARD RATIO ESTIMATES AND 95% CONFIDENCE INTERVALS FOR DEATH USING DIFFERENT SIZES OF SUBCOHORTS ($N=13,658$)

	<i>AIDS at treatment initiation</i>			
	<i>23 cohorts ≈500 per cohort</i>	<i>65 cohorts ≈200 per cohort</i>	<i>134 cohorts ≈100 per cohort</i>	<i>271 cohorts ≈50 per cohort</i>
Fixed effects	1.99 (1.74–2.28)	2.00 (1.75–2.28)	2.00 (1.75–2.29)	2.03 (1.78–2.33)
Random effects ^a	1.97 (1.72–2.25)	1.94 (1.70–2.21)	1.88 (1.65–2.14)	1.84 (1.62–2.10)
Stratified	2.00 (1.75–2.28)	2.00 (1.75–2.28)	1.99 (1.74–2.28)	2.02 (1.77–2.32)

^aModel assumes a Gaussian random effects distribution.

All analyses were adjusted for age, sex, nadir CD4 count, year of initiation, initial ART regimen, and AIDS defining condition at ART initiation.

death among patients from several cohorts throughout Latin America and the Caribbean. Given the substantial differences between cohorts in mortality rates and patterns (e.g., Fig. 1), large differences in the rates of AIDS at ART initiation, and heterogeneity across cohorts in patient follow-up, ascertainment of death, and the definition of AIDS (see Materials and Methods), it seemed plausible that the association between AIDS at ART initiation and mortality would substantially differ based on the analysis approach employed. Although there were certainly differences in estimates from the naive approach that ignored patient cohort, we were surprised to find little difference between hazard ratio estimates using any of the other different analysis approaches. Even in sensitivity analyses using artificial datasets constructed to augment heterogeneity, hazard ratio estimates were quite comparable between analysis approaches. Although only a case study of a single multicohort dataset, our results suggest that as long as

we account for cross-cohort heterogeneity, the specific approach for addressing heterogeneity may not be critical.

The key distinction between the modeling strategies is their flexibility in accounting for potential differences in baseline hazards and hazard ratios between cohorts. In modeling the cohort effect, an investigator can assume the same baseline hazard (naïve and robust marginal approaches), different baseline hazards constrained to be proportional (fixed effects, random effects, and random cohort/treatment effects approaches), or different baseline hazards with no constraints (stratified and meta-analysis approaches). In modeling the hazard ratio, the investigator can assume the same hazard ratio per cohort (naïve and robust marginal, fixed effects, random effects, and stratified approaches) or allow the hazard ratios to vary (random cohort/treatment effects and meta-analysis approaches). Certain approaches to account for heterogeneity may have other desirable properties for specific studies. For

FIG. 3. (A, B) Estimated hazard of death by cohort using a dataset that was modified such that the heterogeneity in baseline hazards between cohorts is more extreme. The *dark line* corresponds to the hazard of death among those without AIDS and the *gray line* corresponds to the hazard of death among those with AIDS ($N=13,658$).

example, the fixed effects approach yields hazard ratios comparing cohorts with a reference cohort, which may be of interest in some studies. However, for many studies, such hazard ratios are not of interest. The meta-analysis approach allows estimation of hazard ratios separately for each cohort and then combines them. However, if the number of events is low, such estimates may be difficult to calculate for some cohorts and may not be interesting.

Previous studies have shown that random effects models perform well in the presence of cohort effects⁹ and cohort-treatment interactions.¹² The results of both our primary and sensitivity analyses are consistent with these findings. We were surprised by the similarity of results in our HIV dataset from the random/fixed effects and stratified models: we were unable to see appreciable differences between estimates even after modifying the dataset to produce what we felt were highly variable baseline hazards. Our sensitivity analyses also confirmed that fixed and random effects models yield different estimates as the numbers of cohorts increase; however, we found that a very large number of cohorts was needed to produce meaningful differences between estimates.

Our study has some limitations. First, our analyses largely ignored loss to follow-up by assuming that censoring was noninformative conditional on study covariates. Although such an assumption is common in practice, it is strong and likely violated in observational HIV studies, and can lead to biased estimates.^{28,29} As seen in the current study and elsewhere,³⁰ LTFU rates can be quite heterogeneous between cohorts, and differential ascertainment of death may explain heterogeneity.³¹ The best approach for handling LTFU is through targeted patient tracing,³² which was not done in this study. Sensitivity analyses can also be employed by assuming specific mortality rates of patients LTFU,^{28,33} but were beyond the scope of this study.

Second, for this study, we assumed proportional hazards for AIDS and other covariates, i.e., that the hazard ratios were constant over time. Formal tests to assess proportional hazards²⁶ showed this assumption was violated for all seven approaches, globally and for the AIDS at treatment initiation covariate (p -value < 0.01). The association between AIDS at ART initiation and death weakened over time. For example, under the stratified approach, the adjusted hazard ratio was 2.26 (95% CI: 1.90–2.70) during the first year, 1.98 (95% CI: 1.36–2.87) during the second year (among those who survived the first year), and (95% CI: 1.06–2.41) during the third year (among those who survived the first 2 years); in subsequent years, the adjusted HR estimates ranged between 0.95 and 1.76. This suggests that, in our setting, we should consider reporting time-specific hazard ratios or use models that do not make such an assumption. The hazard ratios presented in our primary analyses can be thought of as summary measures that underreport the true AIDS hazard ratio shortly after ART initiation and overreport it many years after ART initiation.

Cross-cohort heterogeneity can be caused by many factors, including differences in data collection procedures (e.g., different methods for categorizing AIDS), data quality (e.g., missing data), health system characteristics (e.g., patient transfer procedures), and endpoint ascertainment (e.g., ascertainment of death among those LTFU). We have studied various statistical strategies to account for the heterogeneity between cohorts. However, statistical methods should not be seen as a replacement for continued improvement of data quality.

In general, we prefer more flexible analysis approaches to account for cross-cohort heterogeneity, such as the meta-analysis approach and the stratified approach. We believe that the minor trade-off in precision is worth the reliability in estimates. However, we have reservations about making blanket recommendations for modeling cross-cohort heterogeneity; these results based on seven Latin American and Caribbean cohorts may differ from those in other settings. We encourage investigators to consider both the findings of this study as well as their specific study goals and conditions—background event rates, cohort sizes, heterogeneity of covariate definitions between cohorts—when choosing an approach.

Acknowledgments

This work was supported in part by the National Institutes of Health (NIH), grant NIH 1 UO1 AI069923, and by the Caribbean, Central and South American Network for HIV Epidemiology (CCASAnet) of the International Epidemiologic Databases to Evaluate AIDS Program.

Author Disclosure Statement

No competing financial interests exist.

References

1. Gange SJ, Kitahata MM, Saag MS, *et al.*: Cohort profile: The North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD). *Int J Epidemiol* 2007; 36(2):294–301.
2. Egger M, Ekouevi DK, Williams C, *et al.*: Cohort profile: The international epidemiological databases to evaluate AIDS (IeDEA) in sub-Saharan Africa. *Int J Epidemiol* 2012;41(5):1256–1264.
3. de Wolf F, Sabin C, Kirk O, *et al.*: Developing a multi-disciplinary network for clinical research on HIV infection: The EuroCoord experience. *Clin Invest* 2012;2(3): 255–264.
4. May MT, Ingle SM, Costagliola D, *et al.*: Cohort profile: Antiretroviral Therapy Cohort Collaboration (ART-CC). *Int J Epidemiol* 2014;43(3):691–702.
5. May MT, Hogg RS, Justice AC, *et al.*: Heterogeneity in outcomes of treated HIV-positive patients in Europe and North America: Relation with patient and cohort characteristics. *Int J Epidemiol* 2012;41(6):1807–1820.
6. Klein JP: Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* 1992; 48(3):795–806.
7. Prentice RL and Gloeckler LA: Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 1978;40(8):57–67.
8. Smith-Warner SA, Spiegelman D, Ritz J, *et al.*: Methods for pooling results of epidemiologic studies: The Pooling Project of Prospective Studies of Diet and Cancer. *Am J Epidemiol* 2006;163(11):1053–1064.
9. Glidden DV and Vittinghoff E: Modelling clustered survival data from multicentre clinical trials. *Stat Med* 2004;23(3): 369–388.
10. Katsahian S, Latouche A, Mary JY, *et al.*: Practical methodology of meta-analysis of individual patient data using a survival outcome. *Contemp Clin Trials* 2008;29(2):220–230.
11. O'Quigley J and Stare J: Proportional hazards models with frailties and random effects. *Stat Med* 2002;21(21):3219–3233.

12. Smith CT, Williamson PR, and Marson AG: Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005;24(9):1307–1319.
13. Hoffmann CJ, Fielding KL, Johnston V, *et al.*: Changing predictors of mortality over time from cART start: Implications for care. *J Acquir Immune Defic Syndr* 2011; 58(3):269.
14. Velen K, Lewis JJ, Charalambous S, *et al.*: Comparison of tenofovir, zidovudine, or stavudine as part of first-line antiretroviral therapy in a resource-limited-setting: A cohort study. *PloS One* 2013;8(5):e64459.
15. Engsig FN, Zangerle R, Katsarou O, *et al.*: Long-term mortality in HIV-positive individuals virally suppressed for >3 years with incomplete CD4 recovery. *Clin Infect Dis* 2014;58(9):1312–1321.
16. Phanuphak P, Sirivichayakul S, Jiamsakul A, *et al.*: Transmitted drug resistance and antiretroviral treatment outcomes in non-subtype B HIV-1-infected patients in South East Asia. *J Acquir Immune Defic Syndr* 2014;66(1): 74–79.
17. Low N, Chersich MF, Schmidlin K, *et al.*: Intravaginal practices, bacterial vaginosis, and HIV infection in women: Individual participant data meta-analysis. *PLoS Med* 2011;8(2): e1000416.
18. Heffron R, Donnell D, Rees H, *et al.*: Use of hormonal contraceptives and risk of HIV-1 transmission: A prospective cohort study. *Lancet Infect Dis* 2012;12(1): 19–26.
19. Hasse B, Ledergerber B, Furrer H, *et al.*: Morbidity and aging in HIV-infected persons: The Swiss HIV cohort study. *Clin Infect Dis* 2011;53(11):1130–1139.
20. Calleja JM, Walker N, Cuchi P, *et al.*: Status of the HIV/AIDS epidemic and methods to monitor it in the Latin America and Caribbean region. *AIDS* 2002;16:S3–S12.
21. Tuboi SH, Schechter M, McGowan CC, *et al.*: Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. *J Acquir Immune Defic Syndr* 2009;51(5): 615–623.
22. DerSimonian R and Laird N: Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177–188.
23. Cleveland WS, Grosse E, and Shyu WM: Local regression models. *Stat Models S* 1992:309–376.
24. Schafer JL: *Analysis of Incomplete Multivariate Data*. CRC Press, Boca Raton, FL, 2010.
25. Akaike H: A new look at the statistical model identification. *Automatic control. IEEE Transact* 1974;19(6):716–723.
26. Grambsch PM and Therneau TM: Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81(3):515–526.
27. McGowan CC, Cahn P, Gotuzzo E, *et al.*: Cohort profile: Caribbean, Central and South America Network for HIV research (CCASAnet) collaboration within the International Epidemiologic Databases to Evaluate AIDS (IeDEA) programme. *Int J Epidemiol* 2007;36(5):969–976.
28. Egger M, Spycher BD, Sidle J, *et al.*: Correcting mortality for loss to follow-up: A nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med* 2011;8(1):e1000390.
29. Brinkhof MWG, Pujades-Rodriguez M, and Egger M: Mortality of patients lost to follow-up in antiretroviral treatment programmes in resource-limited settings: Systematic review and meta-analysis. *PloS One* 2009;4(6): e5790.
30. Rosen S, Fox MP, and Gill CJ: Patient retention in antiretroviral therapy programs in sub-Saharan Africa: A systematic review. *PLoS Med* 2007;4(10):e298.
31. Boule A, Schomaker M, May MT, *et al.*: Mortality in patients with HIV-1 infection starting antiretroviral therapy in South Africa, Europe, or North America: A collaborative analysis of prospective studies. *PLoS Med* 2014;11(9): e1001718.
32. Geng EH, Emenyonu N, Bwana MB, *et al.*: Sampling-based approach to determining outcomes of patients lost to follow-up in antiretroviral therapy scale-up programs in Africa. *JAMA* 2008;300(5):506–507.
33. Shepherd BE, Sterling TR, Moore RD, *et al.*: Cross-cohort heterogeneity encountered while validating a model for HIV disease progression among antiretroviral initiators. *J Clin Epidemiol* 2009;62(7):729–737.

Address correspondence to:

Mark Giganti
Department of Biostatistics
Vanderbilt School of Medicine
2525 West End, Suite 11000
Nashville, Tennessee 37203

E-mail: mark.giganti@vanderbilt.edu