MORTALITY AND CANCER RISK IN HIV PATIENTS WITH INCOMPLETE VIRAL SUPPRESSION AFTER ANTIRETROVIRAL THERAPY INITIATION

Jennifer S. Lee

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2017

Approved by:

Stephen R. Cole

Dirk P. Dittmer

Joseph J. Eron Jr.

William C. Miller

David B. Richardson

ABSTRACT

Jennifer S. Lee: Mortality and cancer risk in HIV patients with incomplete viral suppression after antiretroviral therapy initiation
(Under the direction of Stephen R. Cole)

Advances in combination antiretroviral therapy (ART) have led to prolonged survival among people infected with HIV, with clinical focus shifting from acute illnesses to chronic diseases, including malignancies.[1, 2] Effective ART commonly suppresses viral load levels to below the detection limit of assays used in clinical practice in the US, but not all patients on ART are able to achieve virologic suppression to undetectable levels. Detectable HIV RNA under 1,000 copies/mL has been studied as a potential risk factor for increased drug resistance, subsequent virologic failure, and mortality;[3-13] however, viral load measurements in this range are of uncertain clinical significance. HIV patients with low viral load under 1,000 copies/mL may not be receiving optimal clinical management, and the potential adverse consequences of low, detectable HIV RNA, such as the development of cancer and chronic disease, remain unclear. The objective of this project was to examine the clinical significance of low-level detectable HIV RNA under 1,000 copies based on the relationship between a single viral load measurement collected six months after treatment initiation and mortality, and to assess cancer risk among treated HIV patients with low, detectable viral load.

We found that HIV patients with a single low-level viral load measurement between 400 to 999 copies/mL shortly after starting therapy experienced a markedly higher 10-year risk of death (20%) compared to those with viral loads under 20 copies/mL (13%). In fact, these patients faced a similar long-term risk of mortality as patients with very high viral loads

(between 1,000 to 4 million copies/mL) that indicated overt treatment failure (23%). We also found that the risk of a first cancer diagnosis in the 10 years following therapy initiation was 6.9% in our study sample, and did not vary by viral load after controlling for baseline characteristics. Our overall findings highlight the importance of rapid viral load suppression after therapy initiation, and indicate that HIV patients with incomplete viral suppression shortly after starting antiretroviral therapy may require closer clinical monitoring and intervention, such as intensification or change of therapy, in order to increase the prospect of successful treatment response and improved survival.

ACKNOWLEDGEMENTS

First, I thank my advisor and dissertation committee chair, Steve Cole, for investing his time and effort to this project and my training over the past four years. I thank my dissertation committee members—Dirk Dittmer, Joe Eron, Bill Miller, and David Richardson—for sharing their knowledge and providing guidance throughout this process.

I am grateful to the patients, staff, and investigators at the study sites of the Center for AIDS Research Network of Integrated Clinical Systems for their contributions to this project. I thank Chad Achenbach and other CNICS collaborators for their feedback and expertise. Financial support was provided by the National Institutes of Health through the STD/HIV training program (T32 AI007001) and Steve's grant (R01 AI100654).

Finally, I thank my family, friends, and colleagues, near and far, for their support and encouragement along the way. Special thanks to my fellow students in the department for all of our work/therapy sessions, coffee/wine breaks, and gym/burger jaunts. And I thank my master's advisor, Christa Fischer Walker, for telling me years ago to take extra epi courses (even though I really did not want to).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIDS	Acquired immune deficiency syndrome
ART	Antiretroviral therapy
CD4	Cluster of differentiation 4
CFAR	Center for AIDS Research
CI	Confidence interval
CNICS	Center for AIDS Research Network of Integrated Clinical Systems
cpm	Copies per milliliter
DNA	Deoxyribonucleic acid
HIV	Human immunodeficiency virus
HR	Hazard ratio
IDU	Injection drug use
IQR	Interquartile range
INSTI	Integrase strand transfer inhibitor
mL	Milliliter
mm^3	Microliter
MSM	Men who have sex with men
NNRTI	Non-nucleoside reverse transcriptase inhibitor
PI	Protease inhibitor
RD	Risk difference
RNA	Ribonucleic acid
RR	Risk ratio
US	United States

CHAPTER 1. BACKGROUND

Combination antiretroviral therapy suppresses viral replication and improves survival. Of the 37 million people living with HIV globally, 15 million, or 40%, are currently receiving antiretroviral therapy.[14, 15] Untreated HIV infection leads to the development of progressive immunosuppression due to $CD4^+$ T-lymphocyte depletion, resulting in AIDS events and premature death. ART interrupts viral replication, reducing viral load in infected individuals and the risk of HIV transmission, and treatment initiation is currently recommended by the World Health Organization and US Centers for Disease Control and Prevention for all people with HIV, regardless of CD4 count.

Effective ART typically suppresses HIV RNA levels to below the detection limits of assays used in clinical practice for the majority of patients in the US, and has led to an overall decline of average HIV viral load levels in the US[16] as well as decreased incidence of AIDS-defining illnesses.[1] Increased access to effective ART regimens has also led to improved survival in people with HIV, and the life expectancy of treated patients in the United States and Canada has dramatically increased in the past 15 years, approaching that of the general population.[17] As a result, it is estimated that by 2020, over half of all people infected with HIV/AIDS in the US will be over the age of 50.[18]

The burden of cancer is on the rise among HIV patients on treatment. In the US, there are 1.2 million people are living with HIV/AIDS, with an estimated 50,000 new infections every year.[19] As access to ART increases and the life expectancy of people living with

HIV/AIDS continues to normalize in the US, the incidence of age-related conditions, such as cardiovascular disease, diabetes, and osteoporosis, will correspondingly increase.[1] Because age is a primary risk factor for many malignancies, including anal cancer, liver cancer, lung cancer, and Hodgkin lymphoma, cancer has also become an increasingly significant cause of morbidity and mortality among people living with HIV/AIDS who are receiving effective ART. Between 1991 and 2005, the burden of AIDS-defining cancers declined among people living with AIDS, but the burden of all other cancers in people with AIDS increased threefold.[1] Accordingly, the clinical focus of long-term HIV/AIDS care in the US has shifted from the treatment of acute illnesses to the prevention of chronic diseases.[1, 2]

Due to immunosuppression, coinfections with other viruses, and elevated prevalence of certain risk behaviors, the risk of developing some cancers is higher among people infected with HIV compared to the general population.[20, 21] Half of all cancer cases that occur in people living with HIV/AIDS are in excess of expected rates among people who are not infected with HIV.[22] According to a cohort study comparing HIV-positive to HIV-negative individuals, the incidence rate ratio for AIDS-defining cancers was 22.5 and for non-AIDS-defining cancers was 1.9 during the period between 2004 and 2007.[23] Another recent large cohort study indicated that the crude cumulative cancer incidence by age 75 among people with and without HIV was 1.5% and 0.05% for anal cancer, 0.9% and 0.09% for Hodgkin lymphoma, 1.1% and 0.4% for liver cancer, and 3.4% and 2.8% for lung cancer, respectively.[24]

The clinical implications of detectable HIV RNA under 1,000 copies/mL remain unknown. HIV viral load, along with CD4 T-cell count, serves as an indicator of treatment response. Ideally, HIV patients on ART are able to achieve and maintain viral loads that are below the detection limits of commercial assays used in clinical practice. Due to a variety of factors including adverse medication side effects and interactions, treatment type,

regimen complexities, and medical and psychiatric comorbidities, not all HIV patients on ART are able to achieve maximal suppression of viral load to undetectable levels. Only 72% of 9,323 patients from the Antiretroviral Therapy Cohort Collaboration had viral loads of 500 copies/mL or below six months after ART initiation.[25] However, the clinical significance of HIV RNA levels that are in the detectable range but still considered low (typically below 1,000 copies/mL) is unclear.

Previous studies on low, detectable HIV RNA. Low, detectable viral load has been previously studied as a potential risk factor for outcomes including increased drug resistance, subsequent virologic failure, non-Hodgkin lymphoma, and mortality.[3-13, 26] In general, these studies have indicated that low, detectable viremia is associated with adverse outcomes compared to undetectable viral load. People with HIV with persistent low-level viremia (for at least six months) between 50 and 999 copies/mL were found to be at higher risk for virologic failure (defined as >1,000 copies/mL) compared to those with viral load below 50 copies/mL.[5] HIV patients with current, three-month, and six-month lagged viremia of 51–500 copies/mL were found to have an elevated risk of developing non-Hodgkin lymphoma compared to those with viral load at or below 50 copies/mL.[26] Increasing levels of viral load (1–19, 20–399, 400–1,000, etc.) were associated with higher odds of five-year all-cause mortality compared to undetectable viral load, though there was no association after adjusting for CD4 count and other factors.[6]

Low, detectable viral load and cancer. The impact of low, detectable HIV RNA on cancer risk is uncertain and has not been rigorously explored in a longitudinal cohort with consideration of time-varying measurements obtained in clinical practice. Because low HIV RNA may be associated with inflammation[6, 27], it is biologically plausible that low, detectable viral load has predictive value in assessing the long-term risk of developing various cancers, particularly those that may be associated with chronic inflammatory processes or responses to

viral infections. Low, detectable viral load may have a direct oncogenic effect in the development of malignant tumors, or may act through a mechanism of increased chronic inflammation and/or immune dysfunction in tissues with ongoing low-level HIV replication.

No clear definition of low, detectable viral load. Despite previous studies evaluating the effects of low, detectable viral load, its clinical significance remains unclear. Epidemiologic and clinical studies tend to set the upper bound of low viral load somewhere between 200 to 1,000 copies/mL[3-13, 26, 28-30], while laboratory studies generally define low-level viremia as viral loads that are below the detection limit of assays commonly used in clinical settings, which usually range from 20 to 400 copies/mL).[31-33] The latter studies often use ultrasensitive assays that can quantify viral loads down to single copies per milliliter. Because the objective of this project was to characterize low, detectable viral load in a way that was relevant to clinical practice, we focused on low-level viral loads that were below 1,000 copies/mL but above the detection limit of commercial assays.

The varying limits of detection of viral load assays currently used in clinical practice further complicate the issue of evaluating low, detectable viral load. Viral load measurements are subject to an assay's limit of detection, and the limit of detection for viral load assays currently used in clinical settings in the US can range from 20 to 400 copies/mL. The goal of ART is to suppress viral load levels to below assay detection limits, but this can be somewhat arbitrary given assay variability, and it is not clear how to compare undetectable viral load results obtained from assays with different detection limits.

CHAPTER 2. SPECIFIC AIMS

Using longitudinal data from a nationwide cohort of HIV-infected adults engaged in clinical care at eight Center for AIDS Research (CFAR) sites from the CFAR Network of Integrated Clinical Systems (CNICS),[34] we pursued the following specific aims:

Aim 1: Determine whether there is a threshold of detectable HIV viral load under 1,000 copies/mL associated with increased all-cause mortality. We estimated 10-year hazard ratios to ascertain whether there was a threshold of detectable HIV RNA under 1,000 copies/mL six months after ART initiation that was associated with elevated all-cause mortality. We hypothesized that there would be a threshold of viral load under 1,000 copies/mL that corresponded to increased 10-year all-cause mortality compared to viral load under 20 copies/mL.

Aim 2: Assess the impact of detectable HIV viral load under 1,000 copies/mL on first cancer risk. We estimated 10-year risk of first cancer diagnosis among patients with detectable HIV RNA under 1,000 copies/mL six months after ART initiation. We sought to evaluate cancer risk among patients with: (a) viral load under 20 copies/mL; (b) viral load between 20 copies/mL and the threshold identified in Aim 1; (c) viral load between the threshold identified in Aim 1 and 999 copies/mL; and (d) viral load at or above 1,000 copies/mL. In the event that no clear threshold were identified in Aim 1, we planned to use a threshold of 200 copies/mL, as this is currently defined as the cutoff for virologic failure by the

US Department of Health and Human Services and the AIDS Clinical Trials Group.[35] We hypothesized that, among HIV patients with detectable viral load under 1,000 copies/mL six months after ART initiation, those with viral load above the threshold identified in Aim 1 (or 200 copies/mL) would have an elevated 10-year risk of developing cancer compared to those with viral load under 20 copies/mL.

As the number of HIV patients on effective ART continues to rise, patients experiencing detectable HIV RNA under 1,000 copies/mL will be observed more frequently in clinical settings. Therefore, it is important to understand the implications of low, detectable viral load, particularly in relation to mortality risk. Additionally, as advances in ART continue to extend lives, cancer will be of growing public health significance among people living with HIV/AIDS, both in the US and worldwide. This research has the potential to provide important evidence to inform focused cancer intervention and screening practices specifically for treated patients living with HIV/AIDS.

CHAPTER 3. METHODS

3.A. Study design

For this study, we analyzed data from a multisite observational cohort of HIV patients in the US. The CNICS repository maintains standardized demographic, laboratory, medication, diagnosis, health care utilization, and vital status data sourced from electronic medical records and other institutional data systems.[34] Given the increasing life expectancy of HIV patients on treatment and the fact that many cancers occur more frequently in older adults, using extant data on an established patient cohort with long-term follow-up at multiple study sites was practical, efficient, and cost effective compared to enrolling and collecting primary data on a new study population. In addition, because CNICS patients are not recruited outside of routine clinical care, the findings of this study are less subject to volunteer and non-response biases than many observational studies.

3.B. Study population

We used data from the Center for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS), a multicenter clinical cohort of over 30,000 HIV patients in the United States (Figure 3.1). CNICS maintains a clinical data repository from electronic medical record systems to support HIV research.[34] The CNICS cohort includes patients aged 18 years and older who initiated primary care in or after 1995 at one of eight CFAR 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 at Chapel Hill; and University of Washington. CNICS is a dynamic cohort, with approximately 1,400 new patients enrolling and 10% of patients leaving care annually.[34]

Figure 3.1. CNICS study sites.

Source: http://www.uab.edu/cnics/cnics-sites

All participants provided written informed consent to be included in the CNICS cohort, or contributed data with a waiver of written informed consent where approved by local institutional review boards. Upon entry into CNICS, demographic and historical information, including prior diagnoses and antiretroviral treatment, was collected. After enrollment, patient data were prospectively captured at clinic visits and included prescribed medications, laboratory test results, and conditions diagnosed by providers. CNICS participants were typically seen in clinical care every three to four months, though frequency of follow-up was patient specific.

The CNICS cohort was well-suited for this study as it is the largest clinical cohort of HIV patients in the United States; by using data from CNICS, we expected our study findings to be generalizable to treated HIV patients engaged in clinical care in the US. Furthermore, over 30% of the CNICS cohort comprised patients aged 50 years and above (Figure 3.2), ensuring sufficient representation of the population typically at high risk for developing malignancies.

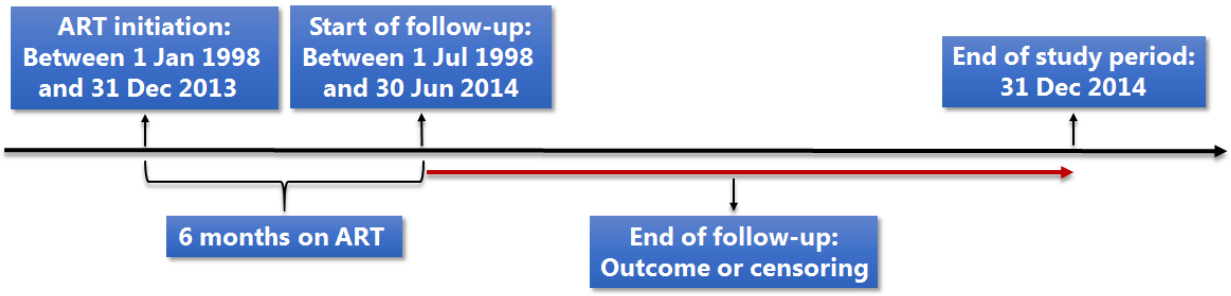
Figure 3.2. CNICS cohort age distribution.

Source: http://www.uab.edu/cnics/data-core/cnics-data-elements

For the study, we examined a subset of the CNICS cohort. The study population included treatment-naïve patients who enrolled in CNICS and initiated combination ART (defined as three or more ART drugs prescribed concurrently) under observation between 1 January 1998 and 31 December 2013, were outcome free six months after ART initiation, and had at least one

recorded viral load measurement six months (-30/+90 days) after ART initiation. For patients who had more than one viral load during the 120-day window, we used the measurement that was collected closest to six months after the date of ART initiation. Date of ART initiation was defined as the date of concurrent prescription of three or more ART drugs. The overall study period was from 1 July 1998 to 31 December 2014 (Figure 3.3).

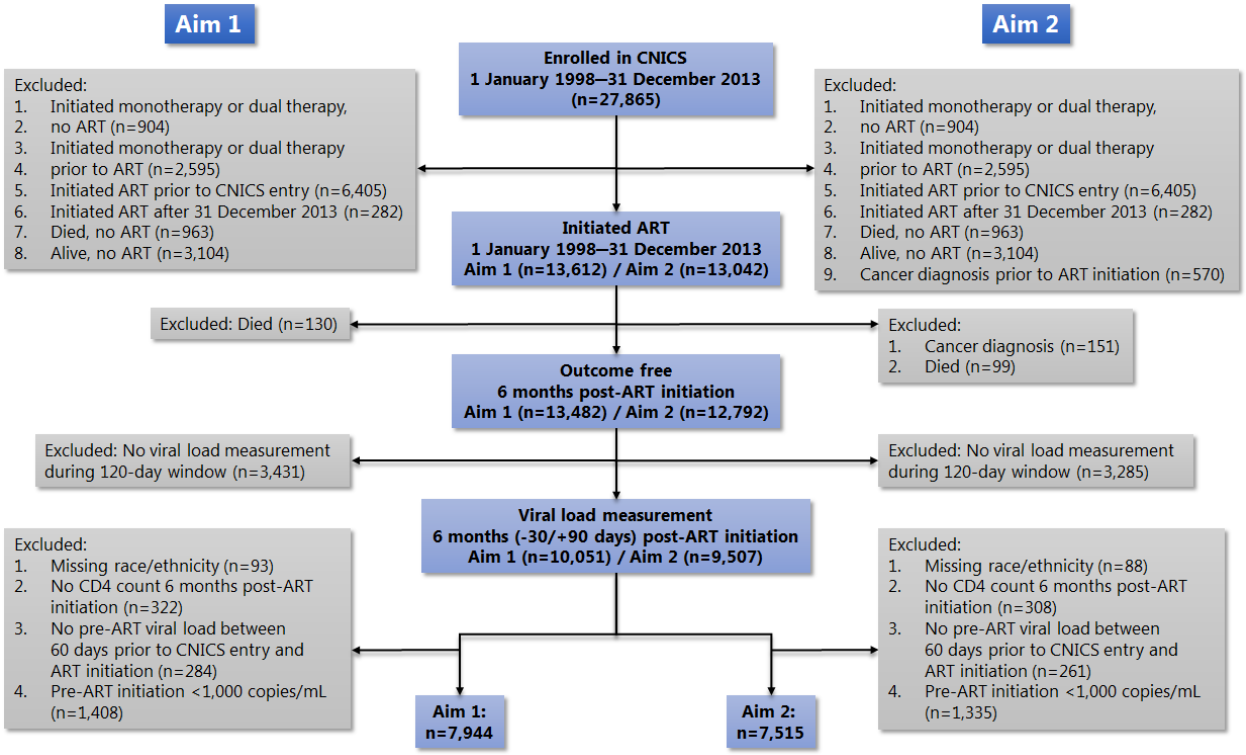
Figure 3.3. Study period (indicated by red arrow).



We excluded patients who initiated monotherapy or dual therapy with no record of initiating combination ART, initiated monotherapy or dual therapy prior to initiating combination ART, or initiated ART over 90 days prior to CNICS entry. Patients with missing race/ethnicity information, no recorded CD4 count six months (-30/+90 days) after ART initiation, no pre-ART viral load measurement recorded between 60 days prior to CNICS entry and 14 days after ART initiation, or pre-ART viral load measurements that suggested unrecorded prior exposure to treatment (<1,000 copies/mL) were also excluded.

There were 27,865 patients who entered the CNICS cohort between 1 January 1998 and 31 December 2013 (Figure 3.4). After applying our eligibility criteria, there were 7,944 patients in the study sample for Aim 1, and 7,515 patients in the study sample for Aim 2.

Figure 3.4. Study eligibility.



3.C. Exposure, outcome, and covariate assessment

For this study, we used extant data collected by CNICS. These data include demographic information, clinical diagnoses, laboratory test results, medications, healthcare utilization, vital status, patient-reported measures and outcomes, antiretroviral resistance mutations, geospatial data, genetic data, and biologic specimens.

Data in CNICS are sourced from point-of-care electronic medical records and other institutional data systems at the eight study sites.[34] Demographic and historical medical information, including prior diagnoses and antiretroviral treatment, are collected on each patient upon enrollment into the CNICS cohort. Once enrolled, laboratory test results, medications, and clinical diagnoses are prospectively captured at outpatient and inpatient

encounters and entered into the electronic medical record by clinicians at CNICS sites. CNICS participants are typically seen in clinical care every three to four months.

Data quality assessments are conducted at all sites prior to data transmission and at the time of submission to the CNICS Data Management Core. After integration into the repository, all data undergo extensive, centralized quality checks, and all data quality issues are reported to CNICS site data managers by the Data Management Core to investigate and correct. Data from each site are updated, reviewed, and integrated into the repository quarterly.

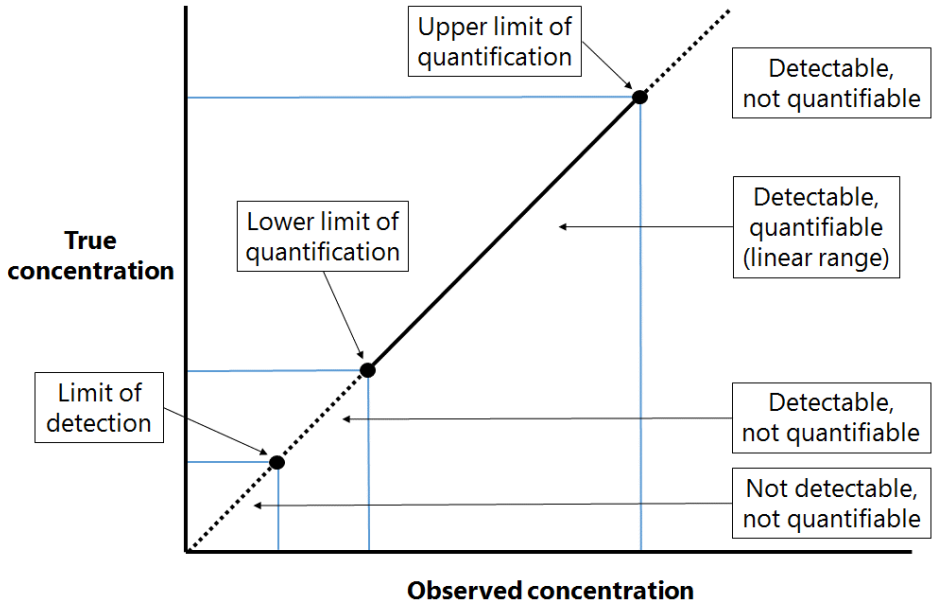
This study was subject to biases inherent in observational studies. Importantly, there may be factors that impact our associations of interest that are not measured in CNICS. That said, the quality of available data in CNICS is high, with the majority of clinical data collected prospectively (the exception being historical data collected at the time of entry into the cohort) and all data undergo extensive, standardized quality assurance procedures.[34]

3.C.1. Exposure: Viral load

The exposure of interest was HIV RNA six months after ART initiation. Viral load measurements were determined by quantitative amplification assays and expressed as the number of HIV copies per milliliter of blood plasma (copies/mL). Viral load assays used in this study varied over time and by CNICS site; the lower limits of detection for the most commonly used assays were 20, 30, 40, 48, 50, 75, and 400 copies/mL. Viral load measurements six months after ART initiation ranged from 6 to over 4 million copies/mL.

Viral load assays have varying lower limits of detection. In assay development, the limit of detection is dependent on the limit of blank, or the highest measurement likely to be observed with a stated probability for a blank, or negative, sample. The limit of blank is estimated by testing replicates of a blank sample and is calculated as the mean blank sample measurement plus the product of the standard deviation of the blank sample measurements and a confidence

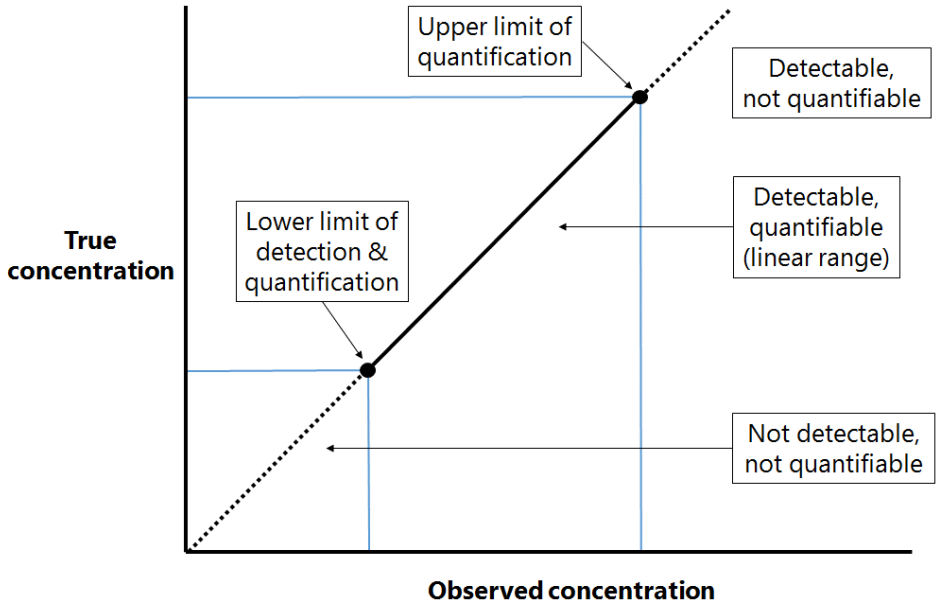
Figure 3.5. Limits of detection and quantification.



level.[36] The limit of detection is the lowest concentration in a sample that can be detected with a stated probability. The limit of detection is estimated by testing replicates of a low concentration sample and is calculated as the limit of blank plus the product of the standard deviation of the low concentration sample measurements and a confidence level.[36, 37] The limit of quantitation/quantification is equivalent to or greater than the limit of detection. The linear range of an assay is bounded by its lower and upper limits of quantification, and represents the region over which the assay provides a linear response with acceptable accuracy (Figure 3.5).

For viral load assays used at the CNICS sites, including the Abbott RealTime HIV-1 Assay, Roche Amplicor HIV-1 Monitor™ Test, and Roche COBAS® AmpliPrep/TaqMan® HIV-1 Test, the limit of detection was equivalent to the lower bound of the assay's linear range (i.e., lower limit of quantification). While labs may specify whether a viral load result that was too low to be quantified was below or above the detection limit, this level of information was not provided in the CNICS data. For this project, the limit of detection was considered equivalent to the lower limit of quantification (Figure 3.6).

Figure 3.6. Limits of detection and quantification (CNICS data).



3.C.2. Aim 1 outcome: All-cause mortality

The outcome of interest for Aim 1 was time to death from any cause. All CNICS sites regularly query the National Death Index and state death certificate records to confirm recorded dates of deaths and capture unrecorded deaths among CNICS patients no longer in care. We used all-cause mortality as our endpoint because cause of death data were unavailable for approximately 35% of deaths recorded in CNICS.

Because vital records are maintained on all patients enrolled in CNICS, including those not currently retained in care, we included deaths among patients no longer being seen in a CNICS clinic in the analyses. We ended the study period on 31 December 2014 to account for reporting delays, which can result in underestimated mortality, and to allow sufficient time for CNICS sites to verify vital records.

3.C.3. Aim 2 outcome: First cancer diagnosis

The outcome of interest for Aim 2 was time to diagnosis of first cancer, excluding nonmelanoma skin cancer. All cancer cases diagnosed through 31 December 2014 and recorded at the CNICS sites were verified by medical record review. Malignancy data collected by CNICS included date of diagnosis, tumor site, diagnosis method (histopathology, clinical exam, radiography, or historical information), histology, stage, and grade. For the analysis, we aggregated all incident cancers into a summary variable.

3.C.4. Covariates

The following variables were included in the analyses as potential confounders:

- *Age*, which was based on year of birth. For all patients, we imputed a birthdate of July 1 of the year of birth, as exact birthdates were not available to protect patient confidentiality.
- *Sex*, which was based on sex at birth. Present sex was used when information about sex at birth was missing.
- *Race/ethnicity*, which was coded as white non-Hispanic, black non-Hispanic, other non-Hispanic, and Hispanic. Separate variables for race and ethnicity were combined to derive this variable.
- *Male-to-male sexual contact*, which was an indicator of ever having had male-to-male sexual contact.
- *Injection drug use*, which was an indicator of ever having injected drugs.
- *Smoking*, which was an indicator of ever having smoked.
- *At-risk alcohol use*, which was an indicator of ever having reported at-risk alcohol use.
- *Pre-ART viral load*, which was viral load collected between 60 days prior to entry in the CNICS cohort and date of ART initiation. If there was no viral load measurement collected during this period, we used viral load measurements collected up to 14 days after ART

initiation, if available. Records of pre-ART viral loads <1,000 copies/mL, which suggested unrecorded prior exposure to treatment, were excluded.

- *Year of ART initiation*, which was based on the date of concurrent prescription of at least three antiretroviral drugs. This date was considered an indicator of starting a combination ART regimen.
- *ART regimen*, which was coded as non-nucleoside reverse transcriptase inhibitor-based, protease inhibitor-based, integrase strand transfer inhibitor-based, or other.
- *CD4+ count*, which serves as a measure of immune function and was expressed as T-lymphocytes per microliter of blood (cells/mm^3).
- *Clinical AIDS diagnosis*, which was an indicator of having been diagnosed with an AIDS-defining illness, according to the 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS.[38]
- *Chronic hepatitis B*, which was based on positive/detectable laboratory test results for hepatitis B surface antigen, DNA, and/or envelope antigen.
- *Chronic hepatitis C*, which was based on positive/detectable laboratory test results for hepatitis C antibody, RNA, and/or genotype.
- *Past cancer diagnosis* (Aim 1 only), which was an indicator of ever having been diagnosed with any cancer, excluding nonmelanoma skin cancer.
- *Statin use*, which was an indicator of ever having used a statin.
- *CNICS site*, which was coded as one of the following: Case Western Reserve University; Fenway Community Health Center, Harvard University; Johns Hopkins University; University of Alabama at Birmingham; University of California, San Diego; University of California, San Francisco; University of North Carolina at Chapel Hill; and University of Washington.

Sex, race/ethnicity, sexual identity, and injection drug use were assessed at entry into the CNICS cohort. Pre-ART viral load, ART regimen, and year of ART initiation were assessed at ART initiation. Age, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, smoking status, at-risk alcohol use, and CNICS site were assessed at study baseline (six months after ART initiation). Restricted quadratic splines were used to model age and CD4 count, with knots at the 5th, 35th, 65th, and 95th percentiles.[39] All other covariates were modeled as indicator variables.

3.D. Statistical analysis

We used SAS version 9.4 (SAS Institute Inc., Cary, NC) for analyses, and R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria) for figures.

3.D.1. Aim 1

Imputing left-censored exposure data. Nearly 70% of viral load observations included in our analyses were reported to be below specified limits of detection. Simple substitution methods (e.g., replacement with the detection limit, half of the detection limit, the detection limit divided by the square root of two, or zero) are typically used to account for left-censored exposure data but can result in substantial bias, particularly when the proportion of censoring is high. Assuming that viral load measurements were left censored at random conditional on observed covariates, an alternate approach is to fit a model with covariates associated with viral load and estimate model parameters by maximum likelihood[40, 41], which produces consistent estimates even when the proportion of left-censoring exceeds 50%.[41] However, this approach is problematic when the data do not closely follow a known parametric distribution.[42]

For this study, we used a nonparametric multiple imputation approach with a left-censoring score model to account for missing data. For each viral load observation, we used logistic regression to estimate the conditional probability of left censoring given all observed covariates and the outcome of interest, death from any cause.[43] We stratified the study cohort into five groups based on quintiles of the predicted probability of being left censored.

Next, we computed nonparametric maximum likelihood estimates of the distribution function of viral load using the Turnbull estimator[44, 45] (equivalent to calculating reverse Kaplan-Meier estimates), stratified by quintiles of the left-censoring score. For each left-censored viral load observation, a random number was generated from a uniform distribution on the interval (0, 1). Within each quintile of the left-censoring score, each random number was matched to the probability distribution function of viral load, and the corresponding viral load was assigned as the imputed viral load value for that left-censored observation. Each imputed viral load value was bounded between zero and the detection limit of the assay that produced the left-censored observation; if an imputed value ended up being above the detection limit, then another random number was generated until the corresponding viral load fell below the detection limit.

We imputed all viral load observations that were too low to be quantified using assays with detection limits of 20, 30, 40, 48, 50, 75, or 400 copies/mL. We note that, with the CNICS data, we were not able to distinguish between left-censored viral loads and viral loads observed at detection limits. One hundred imputed datasets were generated for analysis.

Time-to-event analysis. The start of follow-up for each patient was six months after the date of ART initiation. Patients were followed until death, and data were administratively censored after 10 years or on 31 December 2014. Crude and standardized hazard ratios for 10-

year all-cause mortality from six months post-ART initiation were estimated using the following Cox proportional hazards model[46]:

$$\lambda_k(t) = \lambda_{k,0}(t)\exp\left(\alpha_k x_{k,1} + \beta_k x_{k,2} + \gamma x_3\right)$$

where $x_{k,1}$ is 1 if viral load is between 20 and <k and 0 otherwise, $x_{k,2}$ is 1 if viral load is between k and 999 copies/mL and 0 otherwise, and x_3 is 1 if viral load is above 999 copies/mL and 0 otherwise. The reference category was viral load <20 copies/mL. We estimated hazard ratios for each of the viral load categories, for values of k between 30 and 500 copies/mL. Efron's approximation was used to handle tied event times.[47]

We generated combined point estimates of hazard ratios for 10-year all-cause mortality by averaging across the 100 log hazard ratio estimates from the imputed datasets. We calculated robust standard errors for standardized hazard ratios, and used Rubin's variance estimator[48] to combine variance within and between imputations. These overall hazard ratio and variance estimates were used to construct 95% Wald confidence intervals. The proportional hazards assumption was evaluated by examining plots of the log cumulative hazard by time and testing the product term of viral load and time; no notable violations of this assumption were identified.

Additionally, we computed crude and standardized mortality risk over time using the Nelson-Aalen cumulative hazard function[49, 50], and constructed risk curves[51] averaged across all imputations and stratified by viral load category.

Inverse probability of exposure weights. We used time-fixed inverse probability of exposure weights[52, 53] to control for differences at baseline among patients across the four viral load categories (<20, 20 to <k, k to 999, and >999 copies/mL) and standardize estimates to the total study population. Using a multinomial logistic regression model for each k, we

estimated the conditional probability of being in each viral load category given all observed covariates. To improve efficiency, the weights were stabilized by the marginal probability of being in each viral load category for each k. The stabilized weights had a mean of 1.0 across all values of k for all imputations, with an overall minimum of 0.13 and overall maximum of 15.

For the analysis, we assumed no unmeasured or uncontrolled confounding, no selection bias, no interference, consistency, positive probability of every level of exposure for all strata of covariates, and correct specification of the weight and outcome models.

Alternate analysis. For comparison purposes, we replaced left-censored observations with half of the detection limit, and calculated crude and standardized hazard ratios for 10-year all-cause mortality. In this alternative analysis, the weights used to standardize hazard ratio estimates to the total study population had a mean of 0.96 (range: 0.16, 19) across all k, after truncating at the 99.97 percentile.[52] We also generated crude and standardized risk curves for all-cause mortality over time.

3.D.2. Aim 2

Combining multiple imputation and bootstrap estimation. Multiple imputation is a well-known approach to account for missing data, while the nonparametric bootstrap is often used to estimate standard errors in the absence of a closed-form solution for the standard error of an estimator. However, there is no standard approach to combine both methods. Studies that have combined multiple imputation and bootstrap estimation have generally taken one of two approaches: 1) multiply imputing the original data and then applying the bootstrap to each imputed dataset; or 2) applying the bootstrap to the original data, and then multiply imputing each bootstrap sample. Bootstrap methods are intended to mimic the process of drawing repeated random samples from a population, with the study sample treated

as the population. This seems to point towards using the observed study sample with its original missing data structure intact (i.e., not imputed) when applying the bootstrap. Additionally, there may be efficiency benefits to first bootstrapping the original data and then imputing the bootstrap samples.[54] For this analysis, we chose to take this approach.

First, we drew 200 nonparametric bootstrap samples with replacement from the original study sample. Then, using the nonparametric multiple imputation approach we employed for Aim 1, we imputed all left-censored viral load observations in the original study sample and each of the 200 bootstrap samples. For each viral load observation, we used logistic regression to estimate the conditional probability of left censoring given all observed covariates, first cancer diagnosis, and the competing risk of death. Thirty imputed datasets were generated for the original sample and each bootstrap sample.

To estimate standard errors, we averaged across the point estimates calculated for the 30 imputations for each bootstrap sample, and then computed the standard error of the 200 averaged point estimates. These bootstrap standard error estimates were then used to construct confidence intervals for the combined point estimates calculated from the 30 imputations of the original study sample. Because we first applied the bootstrap to the original data and then multiply imputed the bootstrap samples, it was not necessary to calculate within- and between-imputation variance as an intermediate step.

Time-to-event analysis with competing risks. Studies that characterize cancer risk in HIV patients often measure incidence rates, typically expressed as the number of cancer events per 100,000 person-years, which assume that incident cancers occur at a constant rate over time. Here we estimated the probability of developing cancer over a specific 10-year period, which may provide a more intuitive measure of cumulative cancer risk. Additionally, the majority of previous studies evaluating cancer trends among people with HIV have censored

deaths in their analyses. Failing to use a competing risks approach and treating deaths as censored observations (i.e., using the Kaplan-Meier survival function or standard Cox proportional hazards estimates) ignores the fact that HIV patients may die before being diagnosed with cancer, and will thereby overestimate cancer risk. While censoring competing events may not lead to significant bias when the risk of the competing event is rare, death is not a rare event among HIV patients on treatment. Moreover, censoring competing events may lead to additional bias when the risk of the competing event is differential by exposure, as was the case in this study.

The start of follow-up for each patient was six months after the date of ART initiation. Patients were followed until the earliest of the following: first cancer diagnosis, death, or loss to follow-up (defined as no recorded clinic visit or hospitalization for 18 months). Death from any cause without a cancer diagnosis was considered a competing event. Data were administratively censored after 10 years or on 31 December 2014. We used a proportional subdistribution hazards model[55] to compute nonparametric estimates of the cumulative incidence function of being diagnosed with incident cancer in the presence of the competing risk of death. A minimal amount of random jitter (up to one day) was added to tied event times. We calculated 10-year risk differences and risk ratios and constructed risk curves stratified by viral load category.[51]

To calculate point estimates for cumulative incidence, risk differences, and risk ratios, we averaged across the point estimates calculated from the 30 imputations of the original study sample. We constructed 95% Wald confidence intervals using bootstrap standard error estimates, as described above.

Inverse probability of exposure and censoring weights. We used time-fixed inverse probability of exposure weights[52, 53] to control for differences at baseline among patients across the four viral load categories (<20, 20 to 199, 200 to 999, and >999 copies/mL)

and calculate estimates standardized to the total study population. Using a multinomial logistic regression model, we estimated the conditional probability of being in each viral load category given all observed covariates. To improve efficiency, the exposure weights were stabilized by the marginal probability of being in each viral load category.

Additionally, because we did not have complete outcome ascertainment for Aim 2, we used time-varying inverse probability of censoring weights to account for potentially informative loss to follow-up by viral load category. The distribution of time to loss to follow-up was divided into five intervals. A logistic regression model was used to estimate the conditional probability of remaining in the study cohort during each time interval, given the viral load category and observed covariates. The censoring weights were stabilized by the probability of remaining in the study cohort, conditional on viral load category. The product of the stabilized exposure and censoring weights had a mean of 1.0 (range: 0.12, 13) in the imputed datasets of the original study sample.

For the analysis, we assumed no unmeasured or uncontrolled confounding, no selection bias, no interference, consistency, positive probability of every level of exposure for all strata of covariates, and correct specification of the weight and outcome models.

CHAPTER 4. INCOMPLETE VIRAL SUPPRESSION AND MORTALITY IN HIV PATIENTS ON ANTIRETROVIRAL THERAPY

4.A. Introduction

The goal of antiretroviral therapy (ART) is to restore and maintain the health of people infected with human immunodeficiency virus (HIV) through suppression of HIV replication. In clinical practice, a concentration of HIV ribonucleic acid (RNA) below the detection limits of available assays is considered evidence of viral suppression. Despite advances in ART[56], not all HIV patients on treatment are able to achieve and maintain suppressed viral loads.[57]

Low, detectable HIV RNA under 1,000 copies/mL has been studied as a potential risk factor for drug resistance, virologic failure, cancer, and mortality.[3-13, 26] However, the clinical significance of detectable viral loads in this range remains unclear, despite studies suggesting that patients with HIV or acquired immune deficiency syndrome (AIDS) who have low, detectable viral loads are at higher risk of experiencing adverse health outcomes compared to patients with undetectable viral loads.[3-5, 7, 8, 10, 11, 13, 26]

Uncertainty about the effects of low, detectable viral load is due in part to variability in its definition. Clinical and epidemiologic studies typically set the upper bound of low viral load between 200 and 1,000 copies/mL, while the lower bound tends to be fixed at the detection limit of the viral load assay used in the study.[3-13, 26, 28-30] Defining a range of low, detectable viral load is further complicated when viral load measurements are obtained from multiple assays with different limits of detection, and when a considerable proportion of those measurements fall below varying detection limits.

As access to effective ART scales up and the sensitivity of viral load assays improves over time, the number of HIV patients with low, detectable levels of HIV RNA observed in clinical practice will grow. Therefore, it is important to understand the implications of low, detectable viral load. The objective of this study is to determine whether a clinically significant threshold of detectable viral load under 1,000 copies/mL can be defined based on the relationship between a single viral load measurement collected six months after ART initiation and 10-year mortality.

4.B. Methods

4.B.1. Study population

We used data from the Center for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS), a multicenter clinical cohort that currently includes over 30,000 HIV patients in the United States. CNICS maintains a clinical data repository from electronic medical record systems to support HIV research.[34] The CNICS cohort includes patients aged 18 years and older who initiated primary care in or after 1995 at one of eight CFAR 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 at Chapel Hill; and University of Washington. CNICS is a dynamic cohort, with approximately 1,400 new patients enrolling and 10% of patients leaving care annually.[34]

All participants provided written informed consent to be included in the CNICS cohort, or contributed data with a waiver of written informed consent where approved by local institutional review boards. Upon entry into CNICS, demographic and historical information, including prior diagnoses and antiretroviral treatment, was collected. After enrollment, patient data were prospectively captured at clinic visits and included prescribed medications, laboratory

test results, and conditions diagnosed by providers. CNICS participants were typically seen in clinical care every three to four months, though frequency of follow-up was patient specific.

A total of 27,865 patients entered the CNICS cohort between 1 January 1998 and 31 December 2013. Patients who initiated monotherapy or dual therapy prior to or with no history of starting combination ART (defined as three or more ART drugs prescribed concurrently) (n=3,499), initiated ART prior to entering CNICS (n=6,405), initiated ART after 31 December 2013 (n=282), had no history of initiating ART (n=4,067), died within six months of starting ART (n=130), or did not have at least one viral load measurement six months (-30/+90 days) after ART initiation (n=3,431) were excluded from our study. Patients with missing race/ethnicity information (n=93), no recorded CD4 count six months (-30/+90 days) after ART initiation (n=322), no pre-ART viral load measurement recorded between 60 days prior to CNICS entry and ART initiation (n=284), or pre-ART viral load measurements that suggested unrecorded prior exposure to treatment (<1,000 copies/mL) (n=1,408) were also excluded. A total of 7,944 patients was included in the final study sample.

4.B.2. Mortality ascertainment

The outcome of interest was time to death from any cause. All CNICS sites regularly query the National Death Index and state death certificate records to confirm recorded dates of deaths and capture unrecorded deaths among CNICS patients no longer in care. We used all-cause mortality as our endpoint because cause of death data were unavailable for approximately 35% of deaths recorded in CNICS.

Because vital records are maintained on all patients enrolled in CNICS, including those not currently retained in care, we included deaths among patients no longer being seen in a CNICS clinic in the analyses. We ended the study period on 31 December 2014 to account for

reporting delays, which can result in underestimated mortality, and to allow sufficient time for CNICS sites to verify vital records.

4.B.3. Viral load assessment

The exposure of interest was HIV RNA six months (-30/+90 days) after ART initiation. For patients who had more than one viral load during the 120-day window, we used the measurement that was collected closest to six months after the date of ART initiation. Viral load measurements were determined by quantitative amplification assays and expressed as the number of HIV copies per milliliter of blood plasma (copies/mL). Viral load assays used in this study varied over time and by CNICS site; the lower limits of detection for the most commonly used assays were 20, 30, 40, 48, 50, 75, and 400 copies/mL. Viral load measurements six months after ART initiation ranged from 6 to over 4 million copies/mL.

4.B.4. Statistical analysis

The majority of viral load observations included in our analyses were reported to be below specified limits of detection. We assumed that these viral load measurements were left censored at random conditional on observed covariates, and used a nonparametric multiple imputation approach with a left-censoring score model to account for missing data. For each viral load observation, we used logistic regression to estimate the conditional probability of left censoring given age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking status, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, study site, and death.[43] Restricted quadratic splines were used to model age and CD4 count, with

knots at the 5th, 35th, 65th, and 95th percentiles.[39] We stratified the study cohort into five groups based on quintiles of the predicted probability of being left censored.

Next, we computed nonparametric maximum likelihood estimates of the distribution function of viral load using the Turnbull estimator[44, 45], stratified by quintiles of the left-censoring score, and used these estimates to impute left-censored viral load observations. We imputed all viral load observations that were too low to be quantified using assays with detection limits of 20, 30, 40, 48, 50, 75, or 400 copies/mL, and imputed values were bounded between zero and the assay detection limit. One hundred imputed datasets were generated for analysis.

The start of follow-up for each patient was six months after the date of ART initiation. Patients were followed until death, and data were administratively censored after 10 years or on 31 December 2014. Hazard ratios for 10-year all-cause mortality from six months post-ART initiation were estimated using the following Cox proportional hazards model[46]:

$$\lambda_k(t) = \lambda_{k,0}(t) \exp\left(\alpha_k x_{k,1} + \beta_k x_{k,2} + \gamma x_3\right)$$

where $x_{k,1}$ is 1 if viral load is between 20 and $<k$ and 0 otherwise, $x_{k,2}$ is 1 if viral load is between k and 999 copies/mL and 0 otherwise, and x_3 is 1 if viral load is above 999 copies/mL and 0 otherwise. The reference category was viral load <20 copies/mL. We estimated hazard ratios for each of the viral load categories, for possible threshold values of k between 30 and 500 copies/mL. Efron's approximation was used to handle tied event times.[47]

We used inverse probability of exposure weights[52, 53] to control for differences at baseline among patients across the four viral load categories (<20, 20 to $<k$, k to 999, and >999 copies/mL) and standardize estimates to the total study population. Sex, race/ethnicity, male-to-male sexual contact, and injection drug use were assessed at entry into the CNICS cohort. Pre-ART viral load, year of ART initiation, and ART regimen were assessed at ART initiation.

Age, smoking status, at-risk alcohol use, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site were assessed at study baseline. Restricted quadratic splines were used to model age and CD4 count. Using a multinomial logistic regression model for each k, we estimated the conditional probability of being in each viral load category, and calculated stabilized weights. The weights had a mean of 1.0 across all k for all imputations, with an overall minimum of 0.13 and overall maximum of 15.

We generated combined point estimates of hazard ratios for 10-year all-cause mortality by averaging across the 100 log hazard ratio estimates from the imputed datasets. We calculated robust standard errors for standardized hazard ratios, and used Rubin's variance estimator[48] to combine variance within and between imputations. These overall hazard ratio and variance estimates were used to construct 95% Wald confidence intervals. The proportional hazards assumption was evaluated by examining plots of the log cumulative hazard by time and testing the product term of viral load and time; no notable violations of this assumption were identified. We computed mortality risk over time using the Nelson-Aalen cumulative hazard function[49, 50], and constructed risk curves[51] averaged across all imputations and stratified by viral load category.

Additionally, for comparison, we replaced left-censored observations with half of the detection limit, and calculated hazard ratios for 10-year all-cause mortality and generated risk curves. In this alternate analysis, the weights used to standardize hazard ratio estimates to the total study population had a mean of 0.96 (range: 0.16, 19) across all k, after truncating at the 99.97 percentile.[52] We used SAS version 9.4 (SAS Institute Inc., Cary, NC) for analyses, and R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria) for figures.

4.C. Results

We identified 7,944 CNICS patients (49,118 person-years) who met our study inclusion criteria (Table 4.1). Of the patients included in the study, the median age at baseline (six months after ART initiation) was 40 (interquartile range [IQR]: 32, 46) years, 83% were male, 45% were white/Caucasian, 37% were black/African American, 62% identified as men who have sex with men, and 12% reported having ever injected drugs. The median pre-ART viral load was 74,827 (IQR: 22,394, 237,780) copies/mL, median year of ART initiation was 2007 (IQR: 2003, 2010), and median CD4 count was 349 (IQR: 193, 532) cells/mm^3. At baseline, 29% of study patients had been diagnosed with AIDS, 47% were prescribed a non-nucleoside reverse transcriptase inhibitor-based regimen, and 40% had been prescribed a protease inhibitor-based regimen. Patients were followed for a median of 6.2 (IQR: 3.5, 10) years, and 862 deaths from any cause were recorded during the study period.

Of the patients included in this study, 68% had viral loads six months after ART initiation that were left censored at assay detection limits (Appendix 4.1). After imputation, an average of 57% of all viral load measurements (84% of left-censored observations) fell below 20 copies/mL. Fifteen percent of study patients had viral loads of at least 1,000 copies/mL six months after ART initiation. Plots of the distribution of viral load comparing nonparametric multiple imputation to simple substitution (replacing left-censored observations with half of the detection limit) are shown in Figure 4.1.

Standardized hazard ratio estimates at specific values of k are shown in Table 4.2 (see Appendix 4.2a for crude hazard ratio estimates). Viral load measurements were categorized as <20 copies/mL (reference group), 20 to <k copies/mL, k to 999 copies/mL, and >999 copies/mL. In aggregate, viral load measurements of 20 to 999 copies/mL were not associated with increased 10-year all-cause mortality, when compared to viral loads under 20 copies/mL (standardized hazard ratio [HR]: 1.18, 95% confidence interval [CI]: 0.93, 1.50). When

comparing *k* to 999 copies/mL to <20 copies/mL, we observed an increase in the standardized 10-year hazard ratio for mortality at values of *k* discernable at 130 copies/mL (standardized HR: 1.39, 95% CI: 1.02, 1.88). As expected, viral loads >999 copies/mL were strongly associated with increased mortality (standardized HR: 1.96, 95% CI: 1.56, 2.46). Plots of standardized hazard ratio estimates and 95% confidence intervals by *k* indicated there was no demonstrable viral load threshold between 30 and 500 copies/mL associated with a marked increase in 10-year mortality (Figure 4.2; see Appendix 4.2b for plots of crude hazard ratios).

Standardized risk curves for all-cause mortality at specific values of *k* are shown in Figure 4.3 (see Appendix 4.2c for crude risk curves). Again, we observed an increase in the standardized risk of 10-year mortality with increasing viral load at baseline. The average standardized risk of 10-year mortality was approximately 14% among patients with viral loads between 20 and 400 copies, which was similar to the risk among patients with viral loads <20 copies/mL (13%). There was a 20% standardized risk of death among patients with viral loads between 400 and 999 copies/mL, comparable to the risk among patients with viral loads >999 copies/mL (23%).

Hazard ratio estimates for 10-year all-cause mortality and risk curves, generated after replacing left-censored viral load observations with half of the detection limit, are shown in Appendix 4.3a–e. With this simple substitution approach, the majority (58%) of viral load observations were between 20 and 50 copies/mL, while a comparatively low proportion (8%) of viral loads fell below 20 copies/mL. Patterns in hazard ratio estimates were similar to what we observed with estimates calculated using multiply-imputed viral loads, though the magnitude of hazard ratio estimates was higher across all values of *k* using the substitution method of handling left-censored viral load observations. We also observed similar trends in the risk curves for mortality, though there was a notably less steep trajectory in the standardized risk of

mortality over time among patients with viral loads <20 copies/mL based on data generated from simple substitution.

4.D. Discussion

The objective of this study was to determine whether there was a threshold of HIV viral load under 1,000 copies/mL early after the start of therapy associated with increased mortality, while systematically accounting for undetectable viral load results. We did not identify a clear low-level viral load threshold between 30 and 500 copies/mL that corresponded with a marked increase in 10-year all-cause mortality. Rather, we observed a gradual increase in standardized hazard ratio estimates with increasing viral load, discernable at 130 copies/mL. The average standardized 10-year mortality risk among patients with viral loads between 400 and 999 copies/mL at baseline approached the standardized risk of mortality among patients with viral loads between 1,000 and 4 million copies/mL (20% vs. 23%).

The US Department of Health and Human Services and the AIDS Clinical Trials Group currently define virologic failure as one confirmed viral load measurement over 200 copies/mL.[35] Here, using a single measurement after six months of therapy, we observed a 44% increase in the hazard of death among patients with viral loads between 200 and 999, compared to those with viral loads under 20 copies/mL (standardized HR: 1.44, 95% CI: 1.00, 2.07). The average standardized 10-year mortality risk among patients with low-level viral loads between 200 and 999 copies/mL at baseline was 17%, which was higher than the average standardized risk of mortality among patients with viral loads under 20 copies/mL (14%).

In this study, exposure status was based on one viral load measurement collected approximately six months after ART initiation. Because a single detectable viral load measurement could represent either a transient increase in viral load or sustained low-level HIV RNA concentrations in the detectable range, which are likely disparate risk factors, using time-

varying measures of viral load to assess exposure is warranted for future analysis. That said, we observed a clear pattern of increasing 10-year mortality risk with increasing viral load, based on one viral load measurement under 1,000 copies/mL after six months of therapy. We also observed that a single viral load measurement at or above 1,000 copies/mL six months after ART initiation was strongly associated with 10-year mortality. This suggests that a single viral load measurement collected six months after initiating ART remains highly informative regarding the risk of death over 10 years.

By using data from CNICS, the largest clinical cohort of HIV patients in the US, we expect our study findings to be generalizable to treated adult HIV patients engaged in clinical care at academic medical centers in the US. Because patients who died within six months of initiating ART were excluded from the study population, we expect that the results of this study are applicable to patients who start treatment early enough in the disease course to be effective. There may be unmeasured confounding that impacts our findings, and we note that a viral load measurement collected shortly after starting therapy may be a proxy for unmeasured variables, such as socioeconomic status. We assumed that variables included in the analyses were measured without error, which is unlikely for self-reported behaviors such as tobacco, alcohol, and illicit drug use; however, we do not expect measurement error of confounders to be differential by exposure or outcome.

We did not account for adherence, switching, or cessation of ART regimen in the analyses. For each treatment-naïve patient, we considered the first recorded date of concurrent prescription of three or more ART drugs as an indicator of starting a combination ART regimen, and ignored changes in treatment. Approximately 85% of patients included in our study had viral loads under 1,000 copies/mL six months after ART initiation, and we assumed that patients not taking their medication as prescribed were likely assigned to the highest viral load category (>999 copies/mL). We did not evaluate values of k above 500 copies/mL due to the

relatively small number of events among patients with viral loads between 500 and 999 copies/mL six months after starting therapy. We also note that, because we had limited follow-up for patients whose viral loads were assessed with ultrasensitive assays, it would be prudent to reevaluate our estimates after additional person-time has accumulated in the CNICS cohort.

Nearly 70% of viral load observations included in our analyses were left censored. Simple substitution methods (e.g., replacement with a constant value such as the assay detection limit, half the detection limit, the detection limit divided by the square root of two, or zero) are often used to account for left-censored exposure data but can result in substantial bias, particularly when the proportion of censoring is high.[41, 58] Another approach is to impute left-censored exposure data by maximum likelihood[41, 59], but this is problematic when the data do not closely follow a known parametric distribution.[42] Here, we used nonparametric multiple imputation to account for left-censored viral loads. This allowed us to effectively compare undetectable viral load observations collected over time using assays with different detection limits, without having to rely on distributional assumptions. As shown in Figure 4.1, simple substitution resulted in the majority of viral loads being amassed at specific values determined by assay detection limits, while multiple imputation produced a more biologically plausible depiction of the underlying distribution of viral load.

Using simple substitution resulted in far fewer left-censored viral loads that were categorized as under 20 copies/mL (8% vs. 57% with multiple imputation). Standardized hazard ratio estimates indicated an increased hazard of death for patients at all low-level viral loads between 20 and 999 copies/mL at baseline, due to a markedly lower risk of death among patients with viral loads under 20 copies/mL compared to patients in the same viral load category based on multiply-imputed data (4% vs. 13%). Almost 60% of patients with viral loads under 20 copies/mL based on data generated by simple substitution received care from the same study site and initiated ART during a narrow time period, whereas patients in the same

viral load category based on multiply-imputed data represented all study sites and started ART during all years included in the study period. Using simple substitution resulted in violations of positivity, and fitting the weight model to data generated after simple substitution yielded extreme values, which we did not observe when fitting the same weight model to data generated from multiple imputation. Due to these factors, the hazard ratio and risk estimates we calculated using our nonparametric multiple imputation approach were attenuated but likely less biased than estimates calculated using simple substitution data.

Detectable viral loads under 1,000 copies/mL may indicate ongoing low-level HIV replication due to inadequate response to treatment, drug resistance, drug interactions, or incomplete adherence to therapy or care. Occurrences of low, detectable viral load, whatever the underlying cause, will be more commonly observed in HIV patients as access to antiretroviral therapy increases and assay sensitivity improves over time. While we observed an increased hazard of death with low-level viral loads, discernable at 130 copies/mL, this association was largely driven by the elevated mortality risk experienced by patients with viral loads between 400 and 999 copies/mL. Patients with viral loads in this higher range, which suggested partial response to treatment, faced a similar long-term risk of mortality as patients with high viral loads that indicated overt treatment failure. Low-level viral loads between 400 and 999 copies/mL shortly after starting ART appear to place patients at a significantly higher 10-year risk of death than patients with viral loads under 20 copies/mL, and occurrences of viral loads in this range may need to be treated similarly as viral loads that exceed 1,000 copies/mL. Given the importance of rapidly achieving virologic suppression after initiating treatment, further investigation of the causes of unsuppressed viral loads between 400 and 999 copies/mL is warranted.

4.E. Tables and figures

Table 4.1. Demographic, clinical, and behavioral characteristics of 7,944 CNICS patients six months after ART initiation, between 1 July 1998 and 30 June 2014, averaged over 100 imputations.

Characteristic	Total n=7,944	<20 cpm n=4,545	20–999 cpm n=2,184	>999 cpm n=1,215
	No. (%)	No. (%)	No. (%)	No. (%)
Age, years[a]	40 (32, 46)	40 (32, 47)	40 (33, 47)	39 (32, 45)
Male[b]	6,566 (82.7)	3,790 (83.4)	1,833 (83.9)	943 (77.6)
Race/ethnicity[b]				
White, non-Hispanic	3,581 (45.1)	2,163 (47.6)	971 (44.5)	447 (36.8)
Black, non-Hispanic	2,908 (36.6)	1,482 (32.6)	835 (38.2)	591 (48.6)
Other, non-Hispanic	401 (5.1)	254 (5.6)	101 (4.6)	46 (3.8)
Hispanic	1,054 (13.3)	646 (14.2)	277 (12.7)	131 (10.8)
MSM, ever[b]	4,917 (61.9)	2,946 (64.8)	1,331 (60.9)	640 (52.7)
IDU, ever[b]	979 (12.3)	477 (10.5)	277 (12.7)	225 (18.5)
Smoking, ever	2,605 (32.8)	1,451 (31.9)	708 (32.4)	466 (36.7)
At-risk alcohol use, ever	1,197 (15.1)	654 (14.4)	317 (14.5)	226 (18.6)
Pre-ART viral load, cpm[a,c]	74,827 (22,394; 237,780)	56,048 (17,593; 170,000)	111,880 (36,295; 372,848)	93,928 (30,900; 294,966)
Year of ART initiation[a,c]	2007 (2003, 2010)	2008 (2004, 2010)	2006 (2002, 2009)	2004 (2001, 2008)
ART regimen[c]				
NNRTI-based	3,747 (47.2)	2,463 (54.2)	852 (39.0)	432 (35.6)
PI-based	3,188 (40.1)	1,519 (33.4)	1,048 (48.0)	621 (51.1)
INSTI-based	431 (5.4)	313 (6.9)	87 (4.0)	31 (2.6)
Other	578 (7.3)	229 (5.5)	198 (9.1)	131 (10.8)
CD4 count, cells/mm^3 [a]	349 (193, 532)	405 (245, 574)	321 (184, 496)	191 (71, 355)
Clinical AIDS diagnosis	2,313 (29.1)	1,101 (24.2)	740 (33.9)	472 (38.9)
Chronic hepatitis B	209 (2.6)	91 (2.0)	66 (3.0)	52 (4.3)
Chronic hepatitis C	650 (8.2)	327 (7.2)	177 (8.1)	146 (12.0)
Past cancer diagnosis	429 (5.4)	240 (5.3)	121 (5.5)	68 (5.6)
Statin use, ever	250 (3.2)	167 (3.7)	59 (2.7)	24 (2.0)

Abbreviations: AIDS, acquired immunodeficiency syndrome; ART, antiretroviral therapy; INSTI, integrase inhibitor; IQR, interquartile range; NNRTI, non-nucleoside reverse transcriptase inhibitor; PI, protease inhibitor

[a] Median (interquartile range).

[b] Assessed at entry into CNICS cohort.

[c] Assessed at ART initiation.

Figure 4.1. Distribution of viral loads up to 200 copies/mL for 7,944 CNICS patients six months after ART initiation. Dotted line indicates 20 copies/mL. A. After nonparametric multiple imputation of left-censored viral load observations, averaged over 100 imputations; B. After substitution of left-censored viral load observations with half of assay detection limits.

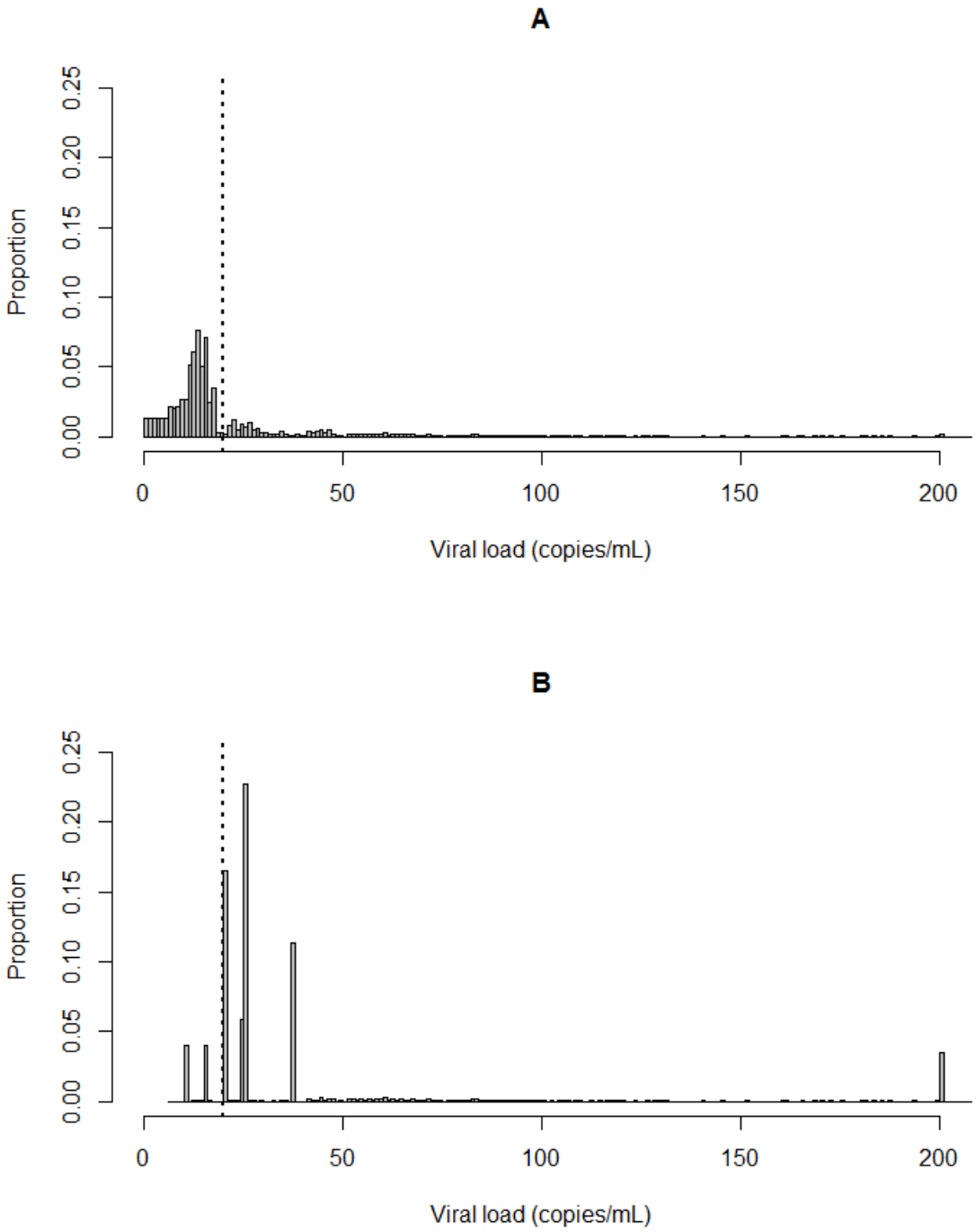


Table 4.2. Standardized hazard ratio estimates for 10-year all-cause mortality for 7,944 CNICS patients, by selected viral load threshold values of *k*, combined from 100 imputations.

	Standardized[a] hazard ratio (95% confidence interval)											
***k*, cpm**	No.[b] of deaths	No.[b] of patients	**<20 cpm**[c]	No.[b] of deaths	No.[b] of patients	**20 to <*k* cpm**	No.[b] of deaths	No.[b] of patients	***k* to 999 cpm**	No. of deaths	No. of patients	**>999 cpm**
20	310	4545	1	—	—	—	230	2184	1.18 (0.93, 1.50)	322	1215	1.96 (1.56, 2.46)[d]
30				54	535	1.17 (0.75, 1.84)	176	1649	1.17 (0.91, 1.50)			
40				75	708	1.19 (0.81, 1.73)	155	1476	1.18 (0.91, 1.52)			
50				96	934	1.17 (0.83, 1.66)	134	1215	1.18 (0.92, 1.53)			
75				121	1273	1.10 (0.79, 1.47)	109	911	1.29 (0.98, 1.70)			
100				142	1454	1.12 (0.85, 1.50)	88	730	1.25 (0.93, 1.68)			
130				152	1592	1.09 (0.83, 1.44)	78	592	1.39 (1.02, 1.88)			
200				174	1776	1.11 (0.85, 1.44)	56	409	1.44 (1.00, 2.07)			
300				189	1912	1.11 (0.86, 1.43)	41	272	1.58 (1.07, 2.35)			
400				197	1991	1.12 (0.87, 1.45)	33	193	1.74 (1.10, 2.74)			
500				201	2041	1.11 (0.86, 1.43)	29	143	1.77 (1.05, 2.99)			

Abbreviation: cpm, copies/mL.

[a] Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site.

[b] Averaged over 100 imputations, rounded to nearest integer.

[c] Viral load <20 copies/mL was reference category across *k*.

[d] Hazard ratio for viral loads >999 copies/mL was unchanged across *k*.

Figure 4.2. Standardized hazard ratios and 95% confidence intervals for 10-year all-cause mortality for 7,944 CNICS patients, by viral load threshold *k*, for viral loads between *k* and 999 copies/mL, combined from 100 imputations. Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site. Upper dashed line indicates hazard ratio for 10-year all-cause mortality for viral loads >999 copies/mL; lower dashed line indicates hazard ratio of 1.

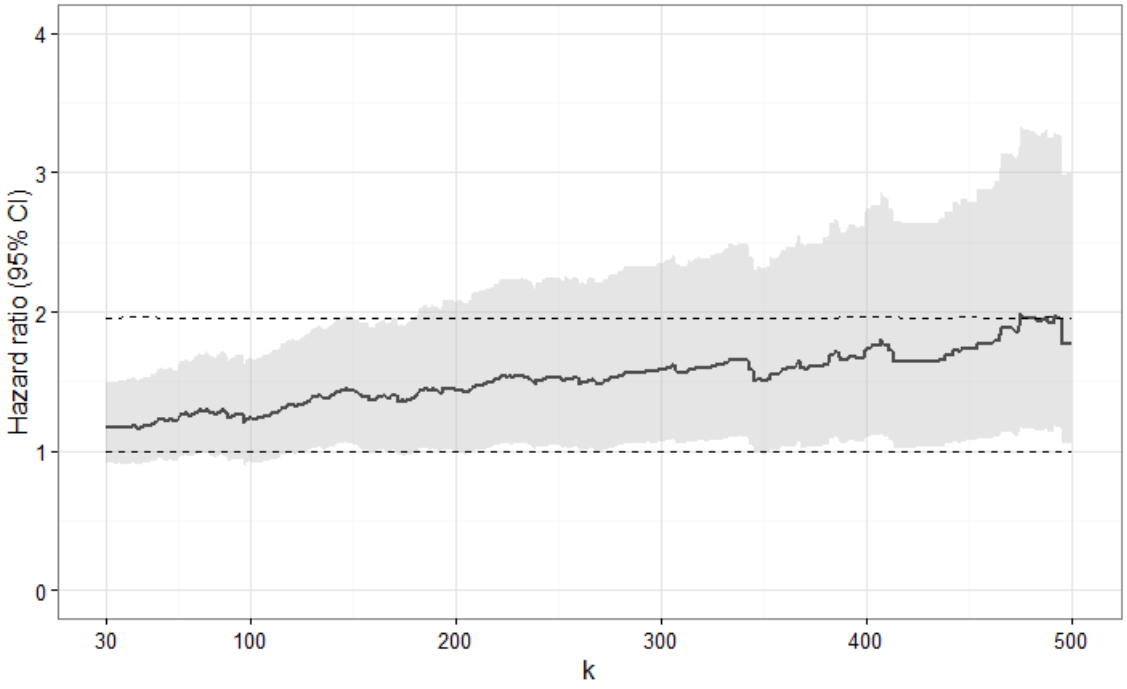
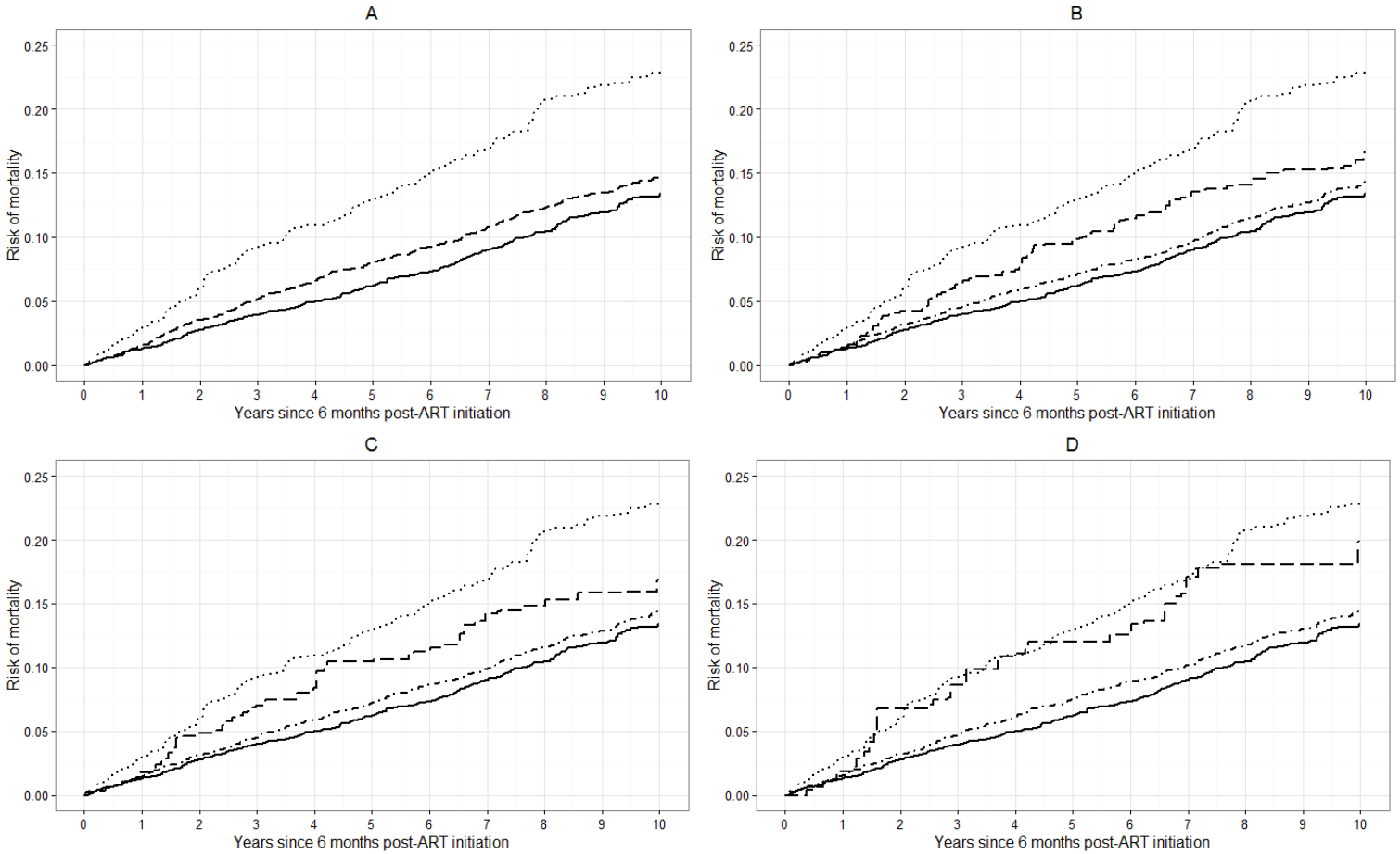


Figure 4.3. Standardized risk curves for all-cause mortality for 7,944 CNICS patients, for selected viral load threshold values of k, stratified by viral load category, averaged over 100 imputations. Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site. Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and k copies/mL, dashed line represents viral loads between k and 999 copies/mL, dotted line represents viral loads >999 copies/mL. A. k = 20 copies/mL; B. k = 130 copies/mL; C. k = 200 copies/mL; D. k = 400 copies/mL.



CHAPTER 5. CANCER RISK IN HIV PATIENTS WITH INCOMPLETE VIRAL SUPPRESSION AFTER INITIATION OF ANTIRETROVIRAL THERAPY

5.A. Introduction

Effective antiretroviral therapy (ART) typically suppresses human immunodeficiency virus (HIV) levels to below the detection limits of assays used in clinical practice in the United States. Treatment with ART has resulted in lower incidence of acquired immune deficiency syndrome (AIDS)-defining illnesses, prolonged survival, and rising incidence of non-AIDS-defining cancers and chronic diseases among people living with HIV.[1] Cancer is the second-leading cause of death in the US[60] and a significant cause of morbidity and mortality in HIV patients.[1, 2] The risk of developing particular cancers may be higher among people infected with HIV compared to the general population due to immunosuppression, oncogenic viral coinfections, and elevated prevalence of certain risk behaviors, such as smoking and alcohol abuse.[20, 21]

Not all HIV patients on treatment are able to achieve and maintain undetectable viral loads, and the impact of low levels of detectable HIV ribonucleic acid (RNA) on the risk of comorbid disease, such as cancer, remains unclear. Prior studies suggest that low-level HIV RNA on ART increases the risk of certain cancers[61, 62]; however, the association between early virologic control and cancer risk has not been evaluated. Optimally, patients initiating ART would achieve undetectable HIV RNA within six months[35], and here we explore whether failure to achieve this milestone is associated with cancer risk. Because low HIV RNA may be associated with ongoing inflammation[6, 27], it is biologically plausible that low, detectable viral load has predictive value in assessing the long-term risk of developing various cancers, particularly those associated with chronic inflammation and viral coinfection.[23, 63] Failure to suppress HIV RNA

after initiation of ART may also be a marker of suboptimal adherence that could influence long-term clinical outcomes. The objective of this study is to examine 10-year cancer risk among HIV patients on ART based on a single low-level viral load measurement collected six months after ART initiation, while accounting for death as a competing risk.

5.B. Methods

5.B.1. Study population

We used data from the Center for AIDS Research (CFAR) Network of Integrated Clinical Systems (CNICS), a multicenter clinical cohort of over 30,000 HIV patients in the United States. CNICS maintains a clinical data repository from electronic medical record systems to support HIV research.[34] The CNICS cohort includes patients aged 18 years and older who initiated primary care in or after January 1995 at one of eight CFAR 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 at Chapel Hill; and University of Washington. CNICS is a dynamic cohort, with approximately 1,400 new patients enrolling and 10% of patients leaving care annually.[34]

All participants provided written informed consent to be included in the CNICS cohort, or contributed data with a waiver of written informed consent where approved by local institutional review boards. Upon entry into CNICS, demographic and historical information, including prior diagnoses and antiretroviral treatment, was collected. After enrollment, patient data were prospectively captured at clinic visits and include prescribed medications, laboratory test results, and conditions diagnosed by providers. CNICS participants were typically seen in clinical care every three to four months, though frequency of follow-up was patient specific.

A total of 27,865 patients entered the CNICS cohort between 1 January 1998 and 31 December 2013. Patients who initiated monotherapy or dual therapy prior to or with no history of starting combination ART (defined as three or more ART drugs prescribed concurrently) (n=3,499), initiated ART prior to entering CNICS (n=6,405), initiated ART after 31 December 2013 (n=282), had no history of initiating ART (n=4,067), were diagnosed with cancer prior to ART initiation (n=570), were diagnosed with cancer within six months of starting ART (n=151), died within six months of starting ART (n=99), or did not have at least one viral load measurement six months (-30/+90 days) after ART initiation (n=3,285) were excluded from our study. Patients with missing race/ethnicity information (n=88), no recorded CD4 count six months (-30/+90 days) after ART initiation (n=308), no pre-ART viral load measurement collected between 60 days prior to CNICS entry and ART initiation (n=261), or pre-ART viral load measurements that suggested unrecorded prior exposure to treatment (<1,000 copies/mL) (n=1,335) were excluded. The final study sample comprised 7,515 patients.

5.B.2. Viral load assessment

The exposure of interest was HIV RNA six months after ART initiation. For patients who had more than one eligible viral load measurement during the 120-day window, we used the measurement that was closest to six months after the date of ART initiation. Viral loads measurements were determined by quantitative amplification assays and expressed as the number of HIV copies per milliliter of blood plasma (copies/mL). Viral load assays used in this study varied over time and by CNICS site; the lower limits of detection for the most commonly used assays were 20, 30, 40, 48, 50, 75, and 400 copies/mL. Observed viral load measurements six months after ART initiation ranged from 6 to over 4 million copies/mL.

5.B.3. Endpoint ascertainment

The outcome of interest was time to diagnosis of first invasive cancer, excluding nonmelanoma skin cancer. All cancer cases diagnosed through 31 December 2014 and recorded at the CNICS sites were verified by medical record review.[64] Cancer data collected by CNICS included date of diagnosis, tumor site, diagnosis method (histopathology, clinical exam, radiography, or historical information), histology, stage, and grade.

Death from any cause was considered a competing risk in the analysis. National Death Index and state death certificate records were queried regularly by all CNICS sites to confirm all recorded dates of deaths.

5.B.4. Statistical analysis

The start of follow-up for each patient was six months after the date of ART initiation. Patients were followed until the earliest of the following: first cancer diagnosis, death, or loss to follow-up (defined as no recorded clinic visit or hospitalization for 18 months). Death from any cause without a cancer diagnosis was considered a competing event. Data were administratively censored after 10 years or on 31 December 2014.

We used a proportional subdistribution hazards model[55] to compute nonparametric estimates of the cumulative incidence function of being diagnosed with incident cancer in the presence of the competing risk of death, and calculated 10-year risk differences and risk ratios and constructed risk curves stratified by viral load category.[51] We drew 200 nonparametric bootstrap samples with replacement from the original study population to estimate standard errors.

The majority of viral load observations included in our analyses were reported to be below specified lower limits of detection (i.e., 20, 30, 40, 48, 50, 75, and 400 copies/mL). For the original study population and each bootstrap sample, we used a nonparametric imputation approach with a censoring score model to account for left-censored viral load data.[65] For each

viral load observation, we used logistic regression to estimate the conditional probability of left censoring given age, sex, race/ethnicity, sexual identity, injection drug use, CD4 count, clinical AIDS status, year of ART initiation, ART regimen, pre-ART viral load, chronic hepatitis status, statin use, smoking status, at-risk alcohol use, CNICS site, death, and incident cancer.[43] Restricted quadratic splines were used to model age and CD4 count, with knots at the 5th, 35th, 65th, and 95th percentiles.[39] We computed nonparametric maximum likelihood estimates[44, 45] of the distribution function of viral load, stratified by quintiles of the predicted probability of being left censored, and used these estimates to impute left-censored viral load observations. Thirty imputed datasets were generated for the original study population and each bootstrap sample.

Patients were assigned to the following exposure categories based on their observed or imputed viral load at baseline (six months after ART initiation): <20 copies/mL, 20 to 199 copies/mL, 200 to 999 copies/mL, and >999 copies/mL. Because the US Department of Health and Human Services and the AIDS Clinical Trials Group currently define virologic failure as one confirmed viral load measurement at or above 200 copies/mL[35], we divided low-level viral loads into two categories, 20 to 199 copies/mL, and 200 to 999 copies/mL. The reference category was viral load <20 copies/mL.

We used inverse probability of exposure weights[52, 53] to control for differences at baseline among patients across the four viral load categories and calculate estimates standardized to the total study population. Sex, race/ethnicity, sexual identity, and injection drug use were assessed at entry into the CNICS cohort. Pre-ART viral load, ART regimen, and year of ART initiation were assessed at ART initiation. Age, CD4 count, clinical AIDS status, chronic hepatitis status, statin use, smoking status, at-risk alcohol use, and CNICS site were assessed at study baseline. Restricted quadratic splines were used to model age and CD4 count. Using a multinomial logistic regression model, we estimated the conditional probability of having a viral load in each viral load category.

Additionally, we estimated inverse probability of censoring weights to account for potentially informative loss to follow-up by viral load category. Both weights were stabilized, and the product of the stabilized weights had a mean of 1.0 in the imputed datasets of the original study sample, with a minimum of 0.12 and maximum of 13. We used SAS version 9.4 (SAS Institute Inc., Cary, NC) for analyses, and R version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria) for figures.

5.C. Results

We identified 7,515 CNICS patients (40,110 person-years) who met our study inclusion criteria (Table 5.1). Of the patients included in the study, the median age at baseline (six months after ART initiation) was 39 (interquartile range [IQR]: 32, 46) years, 82% were male, 45% were white/Caucasian, 37% were black/African American, 61% identified as men who have sex with men, and 13% reported having ever injected drugs. The median pre-ART viral load was 66,691 (IQR: 19,820, 219,732) copies/mL, median year of ART initiation was 2007 (IQR: 2003, 2010), and median CD4 count was 363 (IQR: 207, 541) cells/mm^3. At baseline, 26% of study patients had been diagnosed with AIDS, 41% had been prescribed a protease inhibitor (PI)-based regimen, and 47% were prescribed a non-nucleoside reverse transcriptase inhibitor (NNRTI)-based regimen. Patients were followed for a median of 4.9 (IQR: 2.7, 8.1) years; 14% of the study cohort were followed for 10 years. A total of 290 cancer diagnoses and 560 deaths from any cause without a cancer diagnosis were recorded during the study period, and 1,731 (23%) patients were lost to follow up.

Of the 7,515 patients included in the study, 68% had viral loads six months after ART initiation that were left censored at assay detection limits. After imputation, 56% of all viral loads were under 20 copies/mL, 23% were between 20 and 199 copies/mL, 5% were between 200 and 999 copies/mL, and 15% were over 999 copies/mL. Patients differed across viral load categories at baseline (Table 5.1). Patients with viral loads <20 copies/mL six months after ART

initiation were more likely to be white/Caucasian, identify as men who have sex with men, have higher CD4 counts, have lower pre-ART viral loads, have started ART in the latter half of the study period, have been prescribed an NNRTI-based regimen, and have been prescribed statins. Patients with viral loads <20 copies/mL six months after starting ART were less likely to report injection drug use, have been prescribed a PI-based regimen, have chronic hepatitis, report having ever smoked, or report at-risk alcohol use.

The most common cancers observed in our study population were non-Hodgkin lymphoma (n=39 cases, or 13.4% of all cancer cases), Kaposi sarcoma (n=37; 12.8%), lung cancer (n=30, 10.3%), Hodgkin lymphoma (n=26; 9.0%), prostate cancer (n=22; 7.6%), anal cancer (n=18; 6.8%), breast cancer (n=14; 4.8%), and liver cancer (n=14; 4.8%) (Table 5.2). Crude and standardized 10-year cumulative incidence, risk difference, and risk ratio estimates for first cancer diagnosis are shown in Table 5.3. The crude cancer risk in the study sample was 7.03% (95% confidence interval [CI]: 6.08%, 7.98%). The highest crude cancer risk was observed among patients with viral loads between 200 and 999 copies/mL six months after ART initiation (10.7%), with the risk of cancer diagnosis ranging from 6.60% to 7.67% in the other three viral load categories.

After controlling for baseline characteristics, the overall 10-year cancer risk was 6.90% (95% CI: 5.69%, 8.12%), with little variation in cancer risk by viral load category (range: 6.76% to 7.44%). There was a marked reduction in the cumulative cancer incidence estimate for patients with viral loads between 200 and 999 copies/mL (crude risk of 10.7% vs. standardized risk of 6.82%); race/ethnicity, year of ART initiation, ART regimen, baseline CD4 count, and study site accounted for 75% of the change in estimate in this viral load category. Among patients with viral loads between 200 and 999 copies/mL six months after ART initiation who were diagnosed with cancer, 62% were black (compared to 47% of all cases in the total study population), 52% started ART between 1998 and 2000 (vs. 21%), 63% had been prescribed a PI-based regimen (vs. 46%), and 48% had a CD4 count of less than 200 cells/mm^3 six months after

starting therapy (vs. 36%). Crude and standardized risk curves for 10-year cumulative cancer incidence are shown in Figure 5.1.

The overall standardized risk of death without a cancer diagnosis, which was considered a competing risk in the analysis, was 12.2% (95% CI: 10.2%, 14.2%) (Appendix 5.1–5.2). The risk of death differed by viral load category, ranging from 10.7% among patients with viral loads under 20 copies/mL to 18.1% among patients with viral loads of at least 1,000 copies/mL six months after starting ART.

5.D. Discussion

The objective of this study was to evaluate 10-year cancer risk among HIV patients on antiretroviral therapy with low viral load under 1,000 copies/mL, while accounting for death from any cause as a competing event. The crude 10-year risk of first cancer was highest for patients with viral loads between 200 and 999 copies/mL after six months of therapy, though there was no association between HIV RNA six months after ART initiation and risk of first cancer after controlling for confounders at baseline.

Nearly 70% of the viral load observations used in the analysis fell below assay detection limits, which was expected given that the study population had been on ART for six months at baseline. Prior studies have typically replaced left-censored viral load observations with a constant value, which can result in substantial bias, particularly when the proportion of censoring is high.[41, 58] In a previous study of total mortality in a similar sample of CNICS patients, using a simple substitution approach to account for left-censored viral loads resulted in violations of positivity, unstable weights, and upwardly biased hazard ratio estimates.[65] Here, we used nonparametric multiple imputation to account for left-censored viral loads. This approach allowed for the comparison of undetectable viral load observations collected over time using assays with different detection limits, without imposing assumptions about the underlying distribution of viral load, and likely resulted in less biased estimates.

Studies that characterize cancer risk in HIV patients often measure incidence rates, typically expressed as the number of cancer events per 100,000 person-years, which assume that incident cancers occur at a constant rate over time. Here we estimated the probability of developing cancer over a specific 10-year period, which may provide a more intuitive measure of cumulative cancer risk. Additionally, the majority of previous studies evaluating cancer trends among people with HIV have censored deaths in their analyses. Failing to use a competing risks approach and treating deaths as censored observations (i.e., using the Kaplan-Meier survival function or standard Cox proportional hazards estimates) ignores the fact that HIV patients may die before being diagnosed with cancer, and will thereby overestimate cancer risk. While censoring competing events may not lead to significant bias when the risk of the competing event is rare, we observed nearly twice as many deaths as first cancer diagnoses in our study sample of HIV patients on treatment. Moreover, censoring competing events may lead to additional bias when the risk of the competing event is differential by exposure, as was the case in this study. We expect that we arrived at less biased risk estimates by modeling the cumulative incidence function of cancer, while accounting for death from any cause without a cancer diagnosis as a competing risk.

Here, exposure status was based on a single HIV RNA measurement collected approximately six months after ART initiation, as we considered this a relevant marker of early treatment success. However, a single detectable viral load measurement could represent either a transient increase in viral load or sustained low-level HIV RNA concentrations in the detectable range, so using time-varying measures of viral load to assess exposure is warranted to better understand the dynamic nature of HIV RNA suppression. We did not account for adherence, switching, or cessation of ART regimen in the analyses. For each treatment-naïve patient, we considered the first recorded date of concurrent prescription of three or more ART drugs as an indicator of starting a combination ART regimen, and ignored changes in treatment. We assumed that variables included in the analyses were measured without error, which is unlikely

for self-reported behaviors such as tobacco, alcohol, and illicit drug use; however, we do not expect measurement error of confounders to be differential by exposure or outcome. Outcome misclassification was minimized in this study as cancer cases were confirmed through medical record review, and deaths verified using national and state death records. We also assumed that our models were correctly specified, and that there were no unmeasured or unknown confounders that would significantly impact our findings.

Nearly a quarter of our study sample was lost to follow-up over the study period, and patients were followed for a median of five years. Given that the outcome of interest was 10-year cancer risk, it may be necessary to reassess our risk estimates after additional person-time has accumulated in the CNICS cohort. The precision of our estimates was limited by the relatively small number of cancer cases observed in our study population, so pooling data from other clinical cohorts to verify our results is warranted. Nevertheless, we expect that the results of this study are generalizable to HIV patients receiving care and treatment at academic medical centers in the US, and we observed clinically meaningful trends that highlight potential avenues for cancer screening and prevention among people living with HIV. In this study, non-Hodgkin lymphoma, Kaposi sarcoma, lung cancer, Hodgkin lymphoma, prostate cancer, anal cancer, breast cancer, and liver cancer were the most commonly observed cancer types, consistent with prior studies of cancer in HIV patients after the introduction of ART.[1, 2, 66] This has implications for targeted cancer screening for HIV patients, as well as preventive interventions such as smoking cessation programs and human papillomavirus and hepatitis B vaccination. We also observed a higher proportion of Hodgkin lymphoma cases among patients with lower viral loads, possibly related to immune reconstitution.[67]

We observed a 10-year standardized first cancer risk of 6.9% in our sample of HIV patients after starting therapy. After controlling for baseline characteristics, there was no association between the risk of any first cancer over ten years and early response to ART. This study provides support for existing evidence that cancer continues to pose a significant threat to

HIV patients after ART initiation. While we found that the risk of any first cancer was similar across viral load categories, it would be worth exploring possible differences in cancer types by viral load in future studies of HIV patients on therapy.

5.E. Tables and figures

Table 5.1. Demographic, clinical, and behavioral characteristics of 7,515 CNICS patients six months after ART initiation, averaged over 30 imputations, between 1 July 1998 and 30 June 2014.

Characteristic	Total n=7,515	<20 cpm n=4,281	20–199 cpm n=1,694	200–999 cpm n=393	>999 cpm n=1,147
	No. (%)	No. (%)	No. (%)	No. (%)	No. (%)
Age, years [a]	39 (32, 46)	39 (32, 47)	40 (33, 46)	40 (34, 47)	39 (32, 45)
Male [b]	6,180 (82.2)	3,553 (83.0)	1,421 (82.5)	324 (80.7)	882 (76.9)
Race/ethnicity [b]					
White, non-Hispanic	3,343 (44.5)	2,008 (46.9)	759 (44.8)	160 (40.7)	416 (36.3)
Black, non-Hispanic	2,800 (37.3)	1,426 (33.3)	627 (37.0)	181 (46.0)	566 (49.3)
Other, non-Hispanic	375 (5.0)	238 (5.6)	78 (4.6)	16 (4.2)	43 (3.7)
Hispanic	997 (13.3)	609 (14.2)	230 (13.6)	36 (9.1)	122 (10.6)
MSM, ever [b]	4,611 (61.4)	2,753 (64.3)	1,037 (61.2)	224 (56.9)	597 (52.0)
IDU, ever [b]	938 (12.5)	451 (10.5)	205 (12.1)	67 (17.1)	215 (18.7)
Smoking, ever	2,474 (32.9)	1,366 (31.9)	555 (32.8)	129 (32.8)	424 (37.0)
At-risk alcohol use, ever	1,125 (15.0)	609 (14.2)	244 (14.4)	58 (14.6)	214 (18.7)
Pre-ART viral load, cpm [a,c]	73,320 (22,000; 234,048)	55,133 (17,296; 166,966)	109,225 (35,253; 349,189)	103,356 (35,573; 427,215)	93,071 (30,122; 299,230)
Year of ART initiation [a,c]	2007 (2003, 2010)	2008 (2004, 2010)	2006 (2002, 2009)	2006 (2002, 2009)	2005 (2001, 2008)
ART regimen [c]					
NNRTI-based	3,570 (47.5)	2,335 (54.5)	691 (40.8)	136 (34.6)	409 (35.7)
PI-based	2,997 (39.9)	1,420 (33.2)	782 (46.2)	209 (53.2)	586 (51.1)
INSTI-based	405 (5.4)	290 (6.8)	75 (4.4)	10 (2.5)	30 (2.6)
Other	543 (7.2)	236 (5.5)	147 (8.7)	38 (9.6)	122 (10.6)
CD4 count, cells/mm^3 [a]	356 (201, 537)	412 (254, 581)	338 (197, 511)	299 (177, 456)	204 (75, 362)
Clinical AIDS diagnosis	1,985 (26.4)	918 (21.4)	521 (30.7)	130 (33.1)	417 (36.4)
Chronic hepatitis B	193 (2.6)	82 (1.9)	50 (3.0)	11 (2.9)	49 (4.3)
Chronic hepatitis C	612 (8.1)	307 (7.2)	135 (8.0)	34 (8.8)	135 (11.8)
Statin use, ever	236 (3.1)	158 (3.7)	45 (2.6)	11 (2.8)	22 (1.9)

Abbreviations: AIDS, acquired immunodeficiency syndrome; ART, antiretroviral therapy; cpm, copies per milliliter; INSTI, integrase inhibitor; IDU, injection drug use; IQR, interquartile range; MSM, men who have sex with men; NNRTI, non-nucleoside reverse transcriptase inhibitor; PI, protease inhibitor

[a] Median (interquartile range).

[b] Assessed at entry into CNICS cohort.

[c] Assessed at ART initiation.

Table 5.2. Number (%) of cancers observed in 7,515 CNICS patients, averaged over 30 imputations (rounded to nearest integer).

	Total n=7,515	<20 cpm n=4,281	20–199 cpm n=1,694	200–999 cpm n=393	>999 cpm n=1,147
All cancers	290 (100)	152 (100)	62 (100)	22 (100)	54 (100)
Non-Hodgkin lymphoma	39 (13.4)	17 (11.2)	11 (17.7)	2 (9.1)	9 (16.7)
Kaposi sarcoma	37 (12.8)	17 (11.2)	5 (8.1)	3 (13.6)	12 (22.2)
Lung cancer	30 (10.3)	18 (11.8)	6 (9.7)	1 (4.5)	5 (9.3)
Hodgkin lymphoma	26 (9.0)	18 (11.8)	5 (8.1)	1 (4.5)	2 (3.7)
Prostate cancer	22 (7.6)	12 (7.9)	8 (12.9)	1 (4.5)	1 (1.9)
Anal cancer	18 (6.8)	8 (5.3)	5 (8.1)	3 (13.6)	4 (7.4)
Breast cancer	14 (4.8)	10 (6.6)	3 (4.8)	0	1 (1.9)
Liver cancer	14 (4.8)	7 (4.6)	2 (3.2)	1 (4.5)	4 (7.4)
Skin cancer (melanoma)	11 (3.8)	4 (2.6)	3 (4.8)	1 (4.5)	3 (5.6)
Oral cavity and pharyngeal cancer	10 (3.4)	5 (3.3)	1 (1.6)	1 (4.5)	3 (5.6)
Kidney cancer	8 (2.8)	4 (2.6)	1 (1.6)	2 (9.1)	1 (1.9)
Colon cancer	6 (2.3)	4 (2.6)	2 (3.2)	0	0
Leukemia	6 (2.3)	3 (2.0)	1 (1.6)	0	2 (3.7)
Laryngeal cancer	5 (1.7)	3 (2.0)	2 (3.2)	0	0
Multiple myeloma	5 (1.7)	2 (1.3)	2 (3.2)	0	2 (3.7)
Cervical cancer	4 (1.4)	2 (1.3)	1 (1.6)	0	1 (1.9)
Esophageal cancer	3 (1.0)	2 (1.3)	1 (1.6)	0	0
Thyroid cancer	3 (1.0)	2 (1.3)	0	1 (4.5)	0
Uterine cancer	3 (1.0)	2 (1.3)	1 (1.6)	0	0
Brain and nervous system cancer	2 (0.7)	2 (1.3)	0	0	0
Testicular cancer	2 (0.7)	2 (1.3)	0	0	0
Rectal and rectosigmoid cancer	2 (0.7)	1 (0.7)	1 (1.6)	0	0
Peritoneal & retroperitoneal cancer	2 (0.7)	1 (0.7)	0	0	1 (1.9)
Bladder cancer	1 (0.3)	1 (0.7)	0	0	0
Ovarian cancer	1 (0.3)	1 (0.7)	0	0	0
Soft tissue cancer	1 (0.3)	1 (0.7)	0	0	0
Stomach cancer	1 (0.3)	1 (0.7)	0	0	0
Vaginal cancer	1 (0.3)	0	1 (1.6)	0	0
Vulvar cancer	1 (0.3)	0	0	1 (4.5)	0
Pancreatic cancer	1 (0.3)	0	0	0	1 (1.9)
Small intestine cancer	1 (0.3)	0	0	0	1 (1.9)
Other (unspecified site)	8 (2.8)	3 (2.0)	1 (1.6)	3 (13.6)	1 (1.9)

Abbreviation: cpm, copies per milliliter

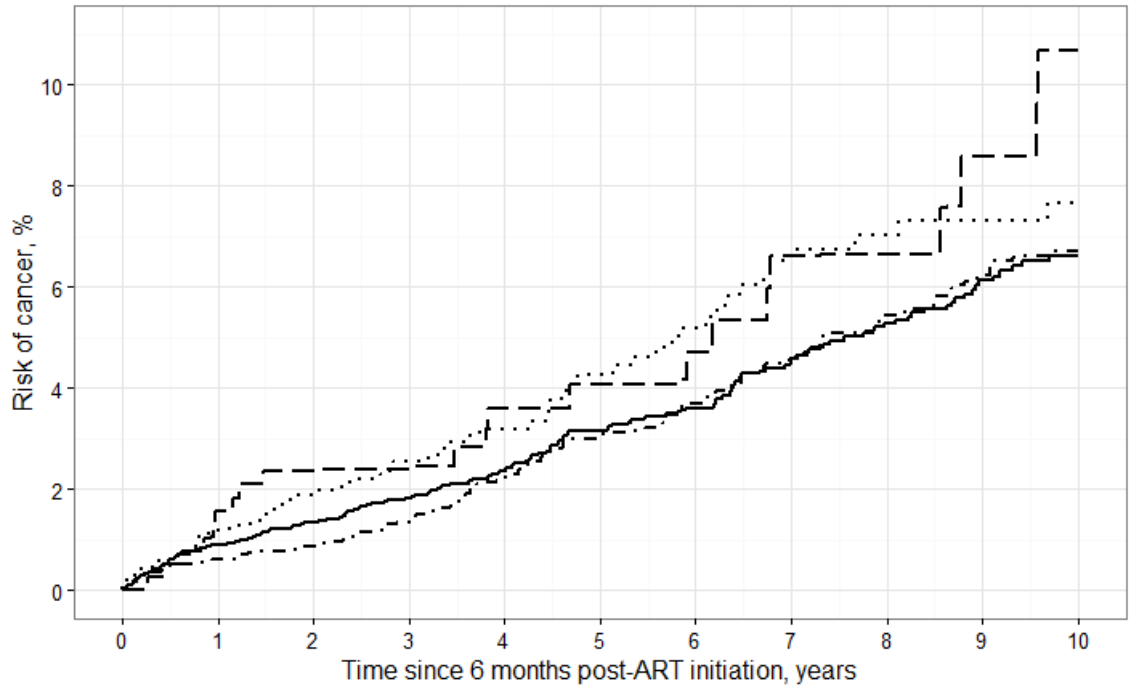
Table 5.3. Crude and standardized 10-year cumulative incidence, risk difference, and risk ratio estimates for first cancer diagnosis in 7,515 CNICS patients, averaged over 30 imputations.

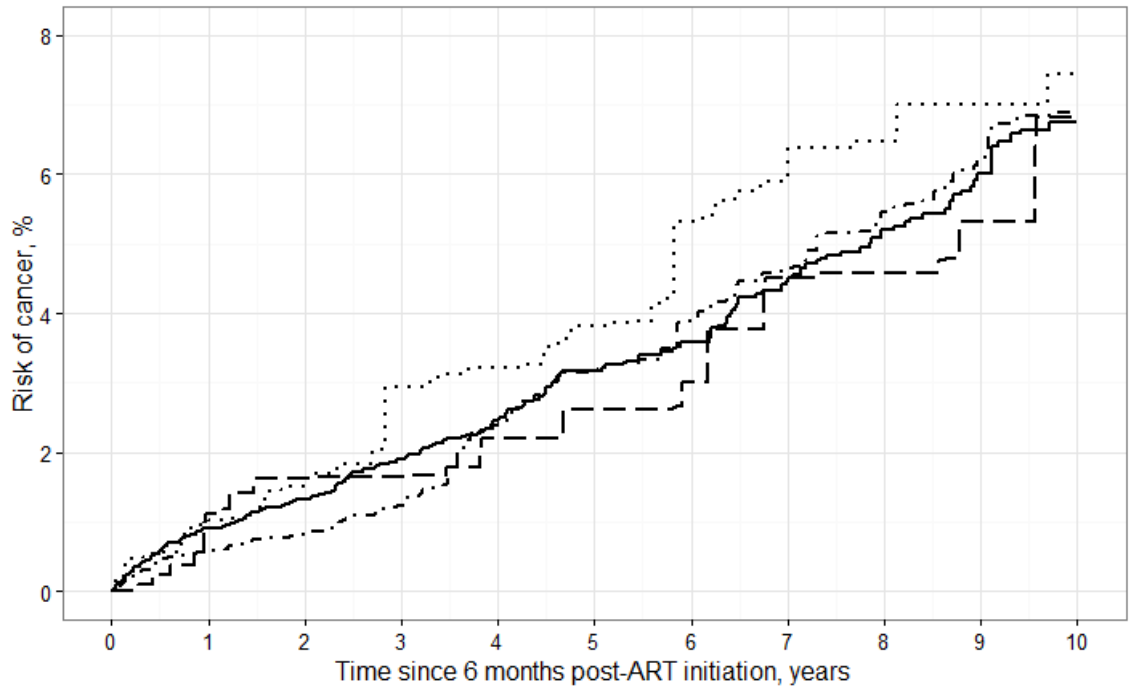
				Crude			Standardized [a]		
	No. of events	No. of patients	Person years	Risk, % (95% CI)	RD, % (95% CI)	RR (95% CI)	Risk, % (95% CI)	RD, % (95% CI)	RR (95% CI)
Total	290	7,515	40,110	7.03 (6.08, 7.98)			6.90 (5.69, 8.12)		
<20 cpm	152	4,281	22,392	6.60 (5.34, 7.86)	0	1	6.76 (5.12, 8.39)	0	1
20 to 199 cpm	62	1,694	9,625	6.71 (5.25, 8.17)	0.10 (-1.74, 1.94)	1.02 (0.73, 1.30)	6.88 (5.08, 8.68)	0.12 (-2.08, 2.33)	1.02 (0.68, 1.36)
200 to 999 cpm	22	393	2,124	10.7 (5.74, 15.6)	4.08 (-0.92, 9.08)	1.62 (0.83, 2.40)	6.82 (3.50, 10.1)	0.06 (-3.73, 3.86)	1.01 (0.43, 1.59)
>999 cpm	54	1,147	5,969	7.67 (5.31, 10.0)	1.07 (-1.69, 3.84)	1.16 (0.72, 1.61)	7.44 (4.10, 10.8)	0.68 (-3.05, 4.41)	1.10 (0.53, 1.67)

Abbreviations: CI, confidence interval; cpm, copies per milliliter; RD, risk difference; RR, risk ratio

[a] Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, chronic hepatitis status), statin use, and study site.

Figure 5.1. Crude and standardized risk curves for first cancer diagnosis in 7,515 CNICS patients, stratified by viral load category, averaged over 100 imputations. Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, chronic hepatitis status, statin use, and study site. Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and 199 copies/mL, dashed line represents viral loads between 200 and 999 copies/mL, dotted line represents viral loads >999 copies/mL.





CHAPTER 6. CONCLUSIONS

6.A. Summary of findings

The objective of Aim 1 of this project was to determine whether there was a threshold of HIV RNA under 1,000 copies/mL early after treatment initiation associated with increased 10-year all-cause mortality. We did not identify a clear low-level viral load threshold between 30 and 500 copies/mL that corresponded with a marked increase in 10-year all-cause mortality. Rather, we observed a gradual increase in standardized hazard ratio estimates with increasing viral load, discernable at 130 copies/mL. The standardized 10-year mortality risk among patients with viral loads between 400 and 999 copies/mL at baseline approached the standardized risk of mortality among patients with viral loads between 1,000 and 4 million copies/mL (20% vs. 23%).

Patients with viral loads between 400 and 999 copies/mL, which suggested partial response to treatment, faced a similar long-term risk of mortality as patients with viral loads between 1,000 and 4 million copies/mL, which indicated overt treatment failure. Low-level viral loads between 400 and 999 copies/mL shortly after starting ART appear to place patients at a significantly higher 10-year risk of death than patients with viral loads under 20 copies/mL, and occurrences of viral loads in this range may need to be treated similarly as viral loads that exceed 1,000 copies/mL.

The objective of Aim 2 was to evaluate the impact of detectable viral load under 1,000 copies/mL on risk of first cancer. In our study sample, we observed a standardized risk of first cancer diagnosis of 6.90% (95% CI: 5.69%, 8.12%). We did not identify an association between the risk of first cancer and viral load, after controlling for baseline characteristics. It is likely that

the relationship between HIV RNA and first cancer risk is time dependent and cannot be adequately captured by a single viral load measurement. The most commonly observed cancers in our overall study population were non-Hodgkin lymphoma, Kaposi sarcoma, lung cancer, Hodgkin lymphoma, prostate cancer, anal cancer, breast cancer, and liver cancer. This provides support for targeted cancer screening for HIV patients, as well as preventive interventions such as smoking cessation programs and human papillomavirus and hepatitis B vaccination.

Nearly 70% of viral load observations included in our analyses were left censored, a significant yet often ignored analytic issue that arises when studying HIV patients on antiretroviral therapy. Here, we used a nonparametric multiple imputation approach to account for left-censored viral loads. Prior studies have typically replaced left-censored viral load observations with a constant value, which can result in substantial bias, particularly when the proportion of censoring is high.[41, 58] Our nonparametric multiple imputation approach allowed us to effectively compare undetectable viral load observations collected over time using assays with different detection limits, without having to rely on distributional assumptions. In our alternate analysis for Aim 1, we showed that simple substitution (replacement with half of assay detection limits) resulted in the majority of viral loads being amassed at specific values determined by assay detection limits, while multiple imputation produced a more biologically plausible depiction of the underlying distribution of viral load. Additionally, using a simple substitution approach to account for left-censored viral loads resulted in violations of positivity, unstable weights, and upwardly biased estimates of association. By using our nonparametric multiple imputation approach, our estimates were attenuated but likely less biased than estimates calculated using simple substitution data.

For Aim 2, we calculated estimates of the cumulative incidence of first cancer diagnosis while accounting for death as a competing event. Studies that characterize cancer risk in HIV patients often measure incidence rates, typically expressed as the number of cancer events per 100,000 person-years, which assume that incident cancers occur at a constant rate over time.

Here we estimated the probability of developing cancer over a specific 10-year period, which may provide a more intuitive measure of cumulative cancer risk. Additionally, the majority of previous studies evaluating cancer trends among people with HIV have censored deaths in their analyses. Failing to use a competing risks approach and treating deaths as censored observations ignores the fact that HIV patients may die before being diagnosed with cancer, and will thereby overestimate cancer risk. While censoring competing events may not lead to significant bias when the risk of the competing event is rare, we observed nearly twice as many deaths as first cancer diagnoses in our study sample of HIV patients on treatment. Moreover, censoring competing events may lead to additional bias when the risk of the competing event is differential by exposure, as was the case in this study. We expect that we arrived at less biased risk estimates by modeling the cumulative incidence function of cancer, while accounting for death from any cause without a cancer diagnosis as a competing risk.

Here, exposure status was based on one viral load measurement collected approximately six months after ART initiation. Because a single detectable viral load measurement could represent either a transient increase in viral load or sustained low-level HIV RNA concentrations in the detectable range, which are likely disparate risk factors, using time-varying measures of viral load to assess exposure is warranted for future analysis. That said, we observed a clear pattern of increasing 10-year mortality risk with increasing viral load, based on one viral load measurement under 1,000 copies/mL after six months of therapy. We also observed that a single viral load measurement at or above 1,000 copies/mL six months after ART initiation was strongly associated with 10-year mortality. This suggests that a single viral load measurement collected six months after initiating ART remains highly informative regarding the risk of death over 10 years. However, a single viral load measurement may be less useful for other long-term outcomes, such as cancer diagnosis.

Because patients who died within six months of initiating ART were excluded from the study population, we expect that the results of this study are generalizable to patients who start

treatment early enough in the disease course to be effective. We assumed that variables included in the analyses were measured without error, which is unlikely for self-reported behaviors such as drug and alcohol use. However, we expect that outcome misclassification was minimized because deaths were verified using national and state death records, and cancer cases were verified by medical record review. We also assumed that our models were correctly specified, and that there were no unmeasured or unknown confounders that would significantly impact our findings.

We did not account for adherence, switching, or cessation of ART regimen in the analyses. For each treatment-naïve patient, we considered the first recorded date of concurrent prescription of three or more ART drugs as an indicator of starting a combination ART regimen, and ignored changes in treatment. Approximately 85% of patients included in our study had viral loads under 1,000 copies six months after ART initiation, and we assumed that patients not taking their medication as prescribed were likely assigned to the highest viral load category (>999 copies/mL).

The precision of our estimates was limited by the relatively small number of events observed in our study population, so pooling data from other clinical cohorts to verify our results is warranted. For Aim 1, we did not evaluate values of k above 500 copies/mL due to the small number of events among patients with viral loads between 500 and 999 copies/mL six months after starting therapy. That said, we expect that the results of this study are likely generalizable to HIV patients receiving care and treatment at academic medical centers in the US. We note that the CNICS cohort is disproportionately male and non-Hispanic white, compared to the overall population of HIV-infected adults in the US.

6.B. Future directions

Using time-varying viral load measurements to assess exposure would better characterize the effect of detectable HIV viral loads under 1,000 copies/mL shortly after starting

therapy on long-term outcomes, such as mortality and cancer risk. Given the high proportion of left-censored viral loads we observed in this study, and the biased estimates that result from using simple substitution to account for these left-censored viral loads, investigating methods to efficiently impute left-censored viral load measurements at multiple time points is warranted.

Viral loads that fall below the detection limits of modern viral load assays generally indicate successful treatment, while high viral loads at or above 1,000 copies/mL after starting therapy likely result from not taking antiretroviral medications as prescribed. Occurrences of detectable viral loads under 1,000 copies/mL may be due to a number of factors, both biological and behavioral. These include inadequate physiologic response to treatment, drug resistance, drug interactions, or incomplete adherence to therapy or care. Given the fact that viral loads in this range will be more commonly observed as access to antiretroviral therapy increases and assay sensitivity improves over time, this study may motivate further research on treatment adherence among CNICS patients, or laboratory research using CNICS biospecimens to investigate inflammatory markers and other factors that may explain partial treatment response.

Furthermore, while it is becoming increasingly evident that the HIV epidemic in the US is shifting to older adults, far less attention and fewer resources have been allocated to examining this trend in the developing world. As access to ART continues to scale up globally, we can expect to observe a similar demographic shift among people living with HIV in developing countries, along with similar increases in the incidence of comorbid chronic conditions. This shift supports further exploration of the prevalence of detectable HIV RNA under 1,000 copies/mL and its potential impact on mortality, cancer, and other chronic diseases among HIV-infected individuals on ART living in resource-limited settings.

6.C. Public health impact

This study highlights the importance of rapid HIV RNA suppression after therapy initiation. Our findings indicate that HIV patients with incomplete viral suppression shortly after starting antiretroviral therapy may require closer clinical monitoring and intervention, such as intensification or change of therapy, in order to increase the prospect of successful treatment response and improved survival.

APPENDIX 4.1. Number (%) of 7,944 CNICS patients with viral load measurements observed or left censored at lower limits of detection for assays most commonly used during study period, by year of start of follow up.

Start of follow-up	All patients	Limit of detection (copies/mL)						
		20	30	40	48	50	75	400
Total	7,944	212 (3.1)	299 (4.4)	1,095 (16.0)	365 (5.3)	1,626 (23.7)	820 (12.0)	238 (3.5)
1998	85	0	0	0	0	32 (37.7)	0	15 (17.7)
1999	268	1 (0.4)	0	1 (0.4)	0	117 (43.7)	0	44 (16.4)
2000	376	0	0	1 (0.3)	0	174 (46.3)	3 (0.8)	39 (10.4)
2001	445	1 (0.2)	0	0	0	191 (42.9)	5 (1.1)	47 (10.6)
2002	384	0	1 (0.3)	1 (0.3)	1 (0.3)	168 (43.8)	26 (6.8)	21 (5.5)
2003	438	0	23 (5.3)	1 (0.2)	0	129 (29.5)	95 (21.7)	21 (4.8)
2004	493	0	47 (9.5)	1 (0.2)	0	146 (29.6)	87 (17.7)	13 (2.6)
2005	451	0	46 (10.2)	2 (0.4)	0	122 (27.1)	112 (24.8)	16 (3.6)
2006	521	0	51 (9.8)	0	0	186 (35.7)	99 (19.0)	15 (2.9)
2007	547	0	50 (9.1)	14 (2.6)	3 (0.6)	199 (36.4)	113 (20.7)	14 (2.6)
2008	624	0	64 (10.3)	66 (10.6)	72 (11.5)	157 (25.2)	94 (15.1)	7 (1.1)
2009	720	0	37 (5.1)	187 (26.0)	146 (20.3)	77 (10.7)	70 (9.7)	4 (0.6)
2010	662	0	0	275 (41.5)	121 (18.3)	48 (7.3)	48 (7.3)	5 (0.8)
2011	665	61 (9.2)	0	288 (43.3)	66 (9.2)	34 (5.1)	50 (7.5)	0
2012	577	123 (21.3)	2 (0.4)	213 (36.9)	26 (4.5)	8 (1.4)	47 (8.2)	1 (0.2)
2013	495	91 (18.4)	0	200 (40.4)	16 (3.2)	6 (1.2)	37 (7.5)	0
2014	193	41 (21.2)	0	63 (32.6)	14 (7.3)	0	19 (9.8)	0

APPENDIX 4.2A. Crude hazard ratio estimates for 10-year all-cause mortality for 7,944 CNICS patients, by selected viral load threshold values of *k*, combined from 100 imputations.

	Crude hazard ratio (95% confidence interval)											
***k*, cpm**	No.[a] of deaths	No.[a] of patients	**<20 cpm**[b]	No.[a] of deaths	No.[a] of patients	**20 to <*k* cpm**	No.[a] of deaths	No.[a] of patients	***k* to 999 Cpm**	No. of deaths	No. of patients	**>999 cpm**
20	310	4545	1	—	—	—	230	2184	1.43 (1.15, 1.78)	322	1251	3.75 (3.18, 4.43)[c]
30				54	535	1.37 (0.89, 2.11)	176	1649	1.45 (1.17, 1.80)			
40				75	708	1.44 (1.00, 2.06)	155	1476	1.43 (1.14, 1.78)			
50				96	934	1.41 (1.01, 1.95)	134	1215	1.45 (1.16, 1.80)			
75				121	1273	1.30 (0.97, 1.73)	109	911	1.61 (1.28, 2.04)			
100				142	1454	1.32 (1.01, 1.73)	88	730	1.64 (1.28, 2.11)			
130				152	1592	1.30 (1.00, 1.68)	78	592	1.80 (1.39, 2.33)			
200				174	1776	1.32 (1.04, 1.69)	56	409	1.90 (1.42, 2.55)			
300				189	1912	1.34 (1.06, 1.70)	41	272	2.06 (1.47, 2.88)			
400				197	1991	1.34 (1.07, 1.69)	33	193	2.33 (1.62, 3.35)			
500				201	2041	1.34 (1.06, 1.68)	29	143	2.77 (1.88, 4.07)			

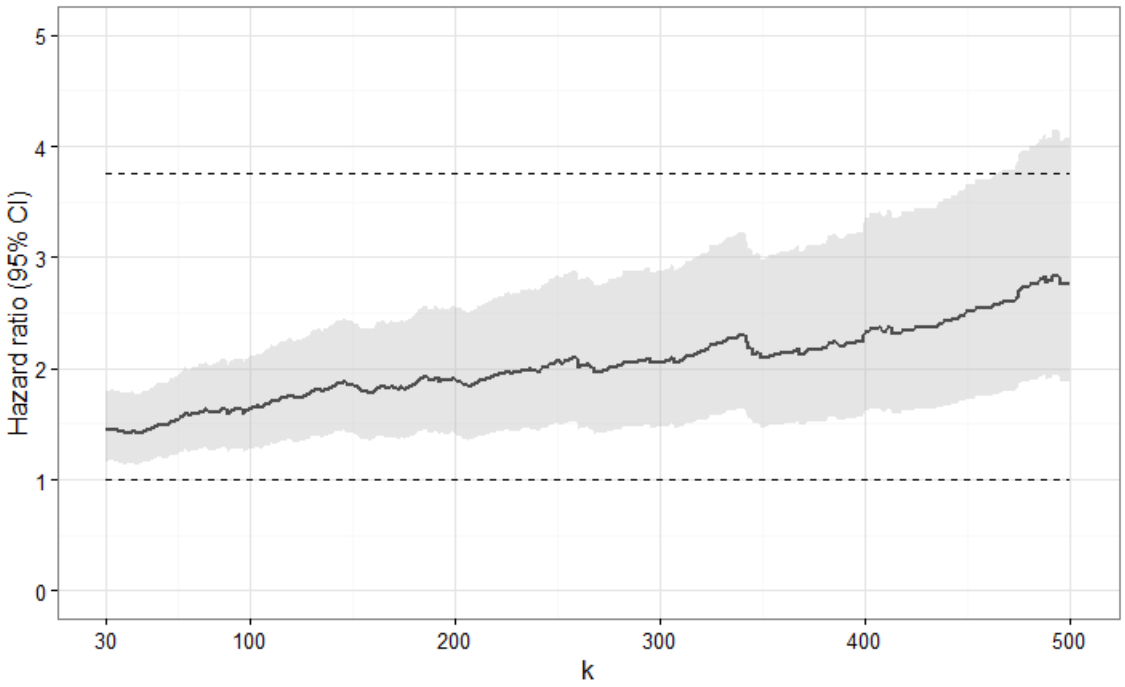
Abbreviation: cpm, copies/mL.

[a] Averaged over 100 imputations, rounded to nearest integer.

[b] Viral load <20 copies/mL was reference category across *k*.

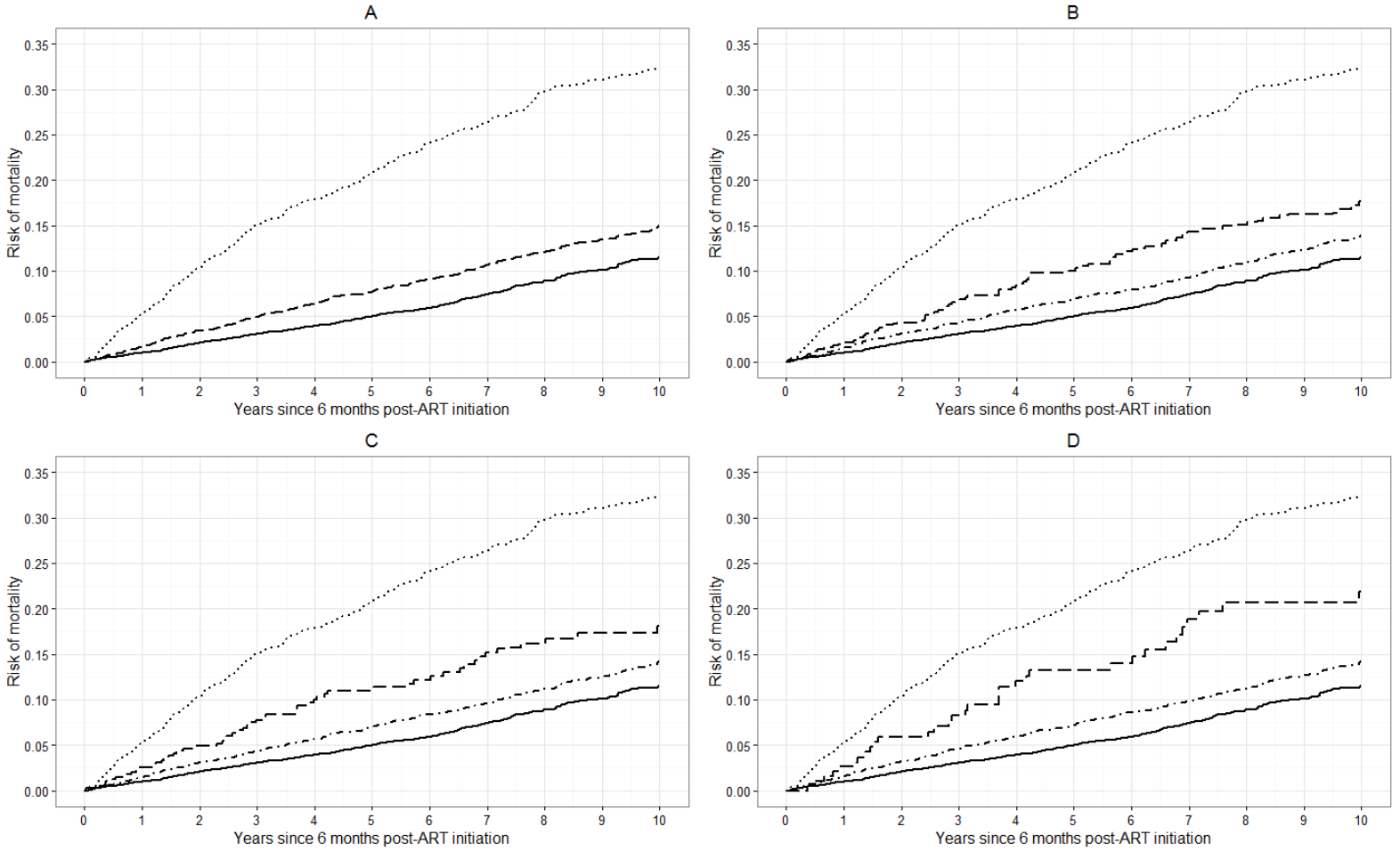
[c] Hazard ratio for viral loads >999 copies/mL was unchanged across *k*.

APPENDIX 4.2B. Crude hazard ratios and 95% confidence intervals for 10-year all-cause mortality for 7,944 CNICS patients, by viral load threshold *k*, for viral loads between *k* and 999 copies/ml, combined from 100 imputations.



Upper dashed line indicates hazard ratio for 10-year all-cause mortality for viral loads >999 copies/mL; lower dashed line indicates hazard ratio of 1.

APPENDIX 4.2C. Crude risk curves for all-cause mortality for 7,944 CNICS patients, for selected viral load threshold values of *k*, stratified by viral load category, averaged over 100 imputations.



Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and *k* copies/mL, dashed line represents viral loads between *k* and 999 copies/mL, dotted line represents viral loads >999 copies/mL. A. *k* = 20 copies/mL; B. *k* = 130 copies/mL; C. *k* = 200 copies/mL; D. *k* = 400 copies/mL.

APPENDIX 4.3A. Crude hazard ratio estimates for 10-year all-cause mortality for 7,944 CNICS patients, by selected viral load threshold values of *k*, with left-censored viral load observations substituted with half of assay detection limit.

	Crude hazard ratio (95% confidence interval)											
***k*, cpm**	No. of deaths	No. of patients	**<20 cpm**[a]	No. of deaths	No. of patients	**20 to <*k* cpm**	No. of deaths	No. of patients	***k* to 999 cpm**	No. of deaths	No. of patients	**>999 cpm**
20	39	674	1	—	—	—	501	6055	1.19 (0.86, 1.65)	322	1251	3.83 (2.75, 5.34)
30				276	3623	1.15 (0.82, 1.61)	225	2432	1.24 (0.88, 1.74)			3.83 (2.75, 5.34)
40				336	4552	1.08 (0.78, 1.51)	165	1503	1.48 (1.05, 2.10)			3.83 (2.75, 5.35)
50				339	4640	1.08 (0.77, 1.50)	162	1415	1.52 (1.07, 2.15)			3.83 (2.75, 5.35)
75				356	4919	1.07 (0.77, 1.48)	145	1136	1.66 (1.16, 2.36)			3.84 (2.75, 5.35)
100				375	5091	1.08 (0.78, 1.50)	126	964	1.69 (1.18, 2.42)			3.83 (2.75, 5.35)
130				384	5221	1.08 (0.78, 1.50)	117	834	1.79 (1.25, 2.58)			3.84 (2.75, 5.35)
200				404	5395	1.09 (0.79, 1.52)	97	660	1.84 (1.27, 2.67)			3.84 (2.75, 5.35)
300				461	5787	1.14 (0.83, 1.59)	40	268	2.10 (1.35, 3.27)			3.83 (2.75, 5.34)
400				468	5862	1.15 (0.83, 1.59)	33	193	2.37 (1.49, 3.78)			3.83 (2.75, 5.34)
500				472	5912	1.15 (0.83, 1.59)	29	143	2.83 (1.75, 4.57)			3.83 (2.75, 5.34)

Abbreviation: cpm, copies/mL.

[a] Viral load <20 copies/mL was reference category across *k*.

APPENDIX 4.3B. Standardized hazard ratio estimates for 10-year all-cause mortality for 7,944 CNICS patients, by selected viral load threshold values of *k*, with left-censored viral load observations substituted with half of assay detection limit.

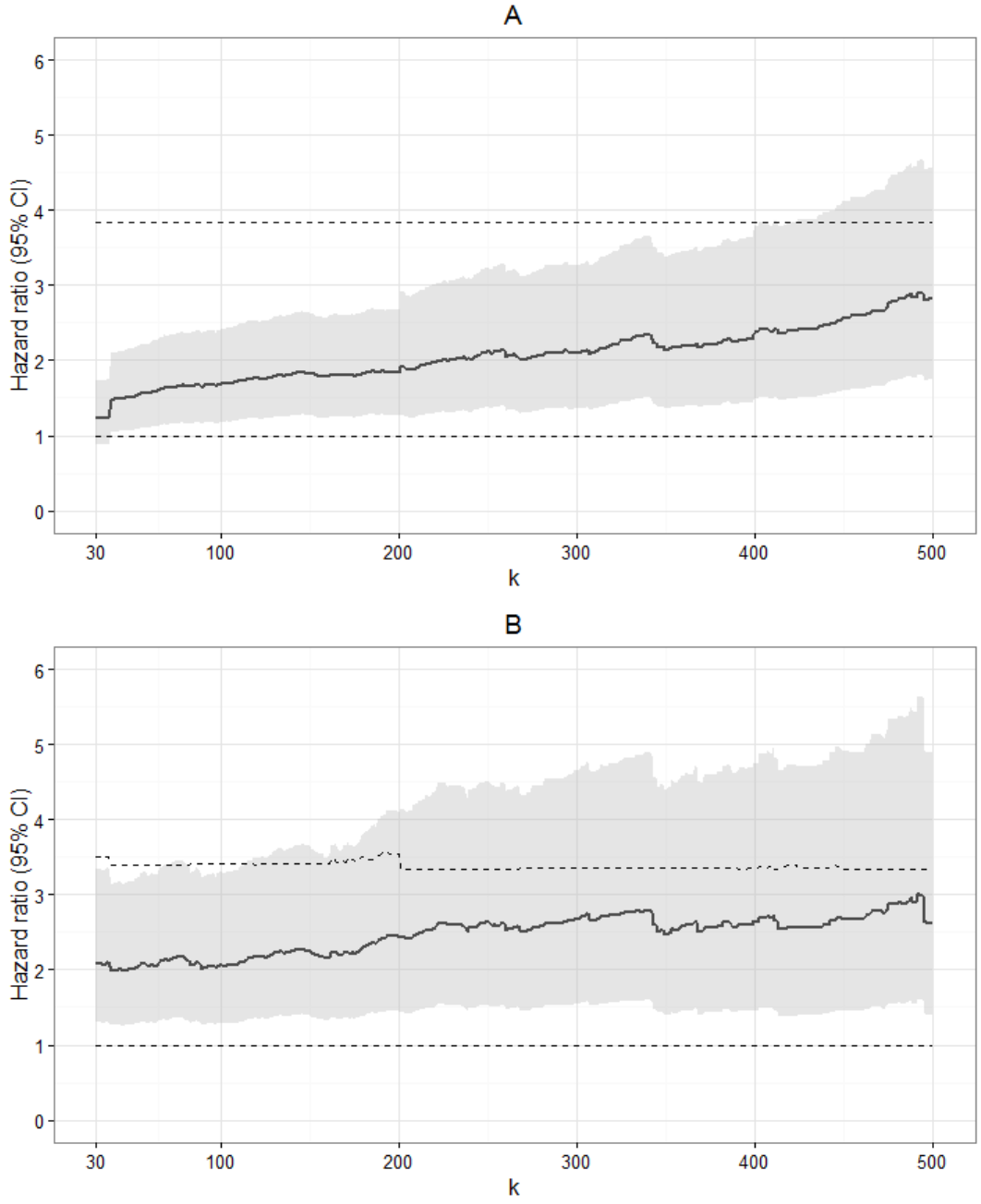
	Standardized[a] hazard ratio (95% confidence interval)											
***k*, cpm**	No. of deaths	No. of patients	**<20 cpm[b]**	No. of deaths	No. of patients	**20 to <*k* cpm**	No. of deaths	No. of patients	***k* to 999 cpm**	No. of deaths	No. of patients	**>999 cpm**
20	39	674	1	—	—	—	501	6055	1.84 (1.19, 2.82)	322	1251	3.34 (2.13, 5.24)
30				276	3623	1.72 (1.07, 2.76)	225	2432	2.09 (1.30, 3.35)			3.51 (2.17, 5.66)
40				336	4552	1.80 (1.16, 2.78)	165	1503	2.00 (1.27, 3.15)			3.39 (2.16, 5.32)
50				339	4640	1.81 (1.17, 2.79)	162	1415	2.01 (1.27, 3.18)			3.40 (2.17, 5.32)
75				356	4919	1.76 (1.14, 2.72)	145	1136	2.17 (1.36, 3.46)			3.40 (2.17, 5.34)
100				375	5091	1.80 (1.17, 2.78)	126	964	2.08 (1.30, 3.33)			3.41 (2.18, 5.35)
130				384	5221	1.78 (1.15, 2.74)	117	834	2.22 (1.38, 3.58)			3.41 (2.17, 5.34)
200				404	5395	1.86 (1.18, 2.92)	97	660	2.44 (1.45, 4.10)			3.54 (2.22, 5.66)
300				461	5787	1.76 (1.15, 2.71)	40	268	2.68 (1.54, 4.65)			3.35 (2.14, 5.25)
400				468	5862	1.79 (1.16, 2.75)	33	193	2.63 (1.46, 4.74)			3.35 (2.14, 5.26)
500				472	5912	1.77 (1.15, 2.73)	29	143	2.63 (1.41, 4.90)			3.35 (2.14, 5.25)

Abbreviation: cpm, copies/mL.

[a] Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site.

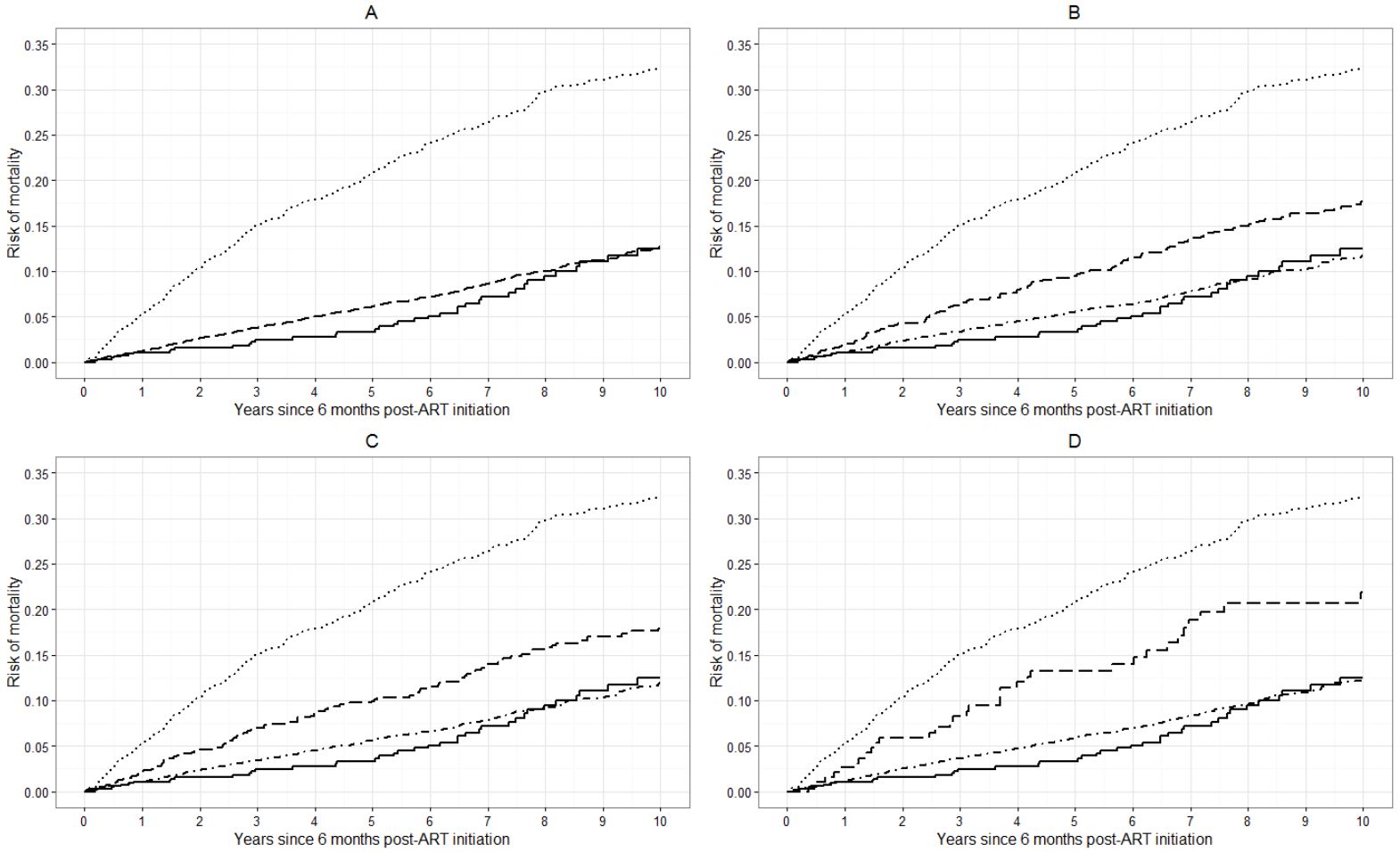
[b] Viral load <20 copies/mL was reference category across *k*.

APPENDIX 4.3C. Crude and standardized hazard ratios and 95% confidence intervals for 10-year all-cause mortality for 7,944 CNICS patients, by viral load threshold *k*, for viral loads between *k* and 999 copies/ml, with left-censored viral load observations substituted with half of assay detection limit.



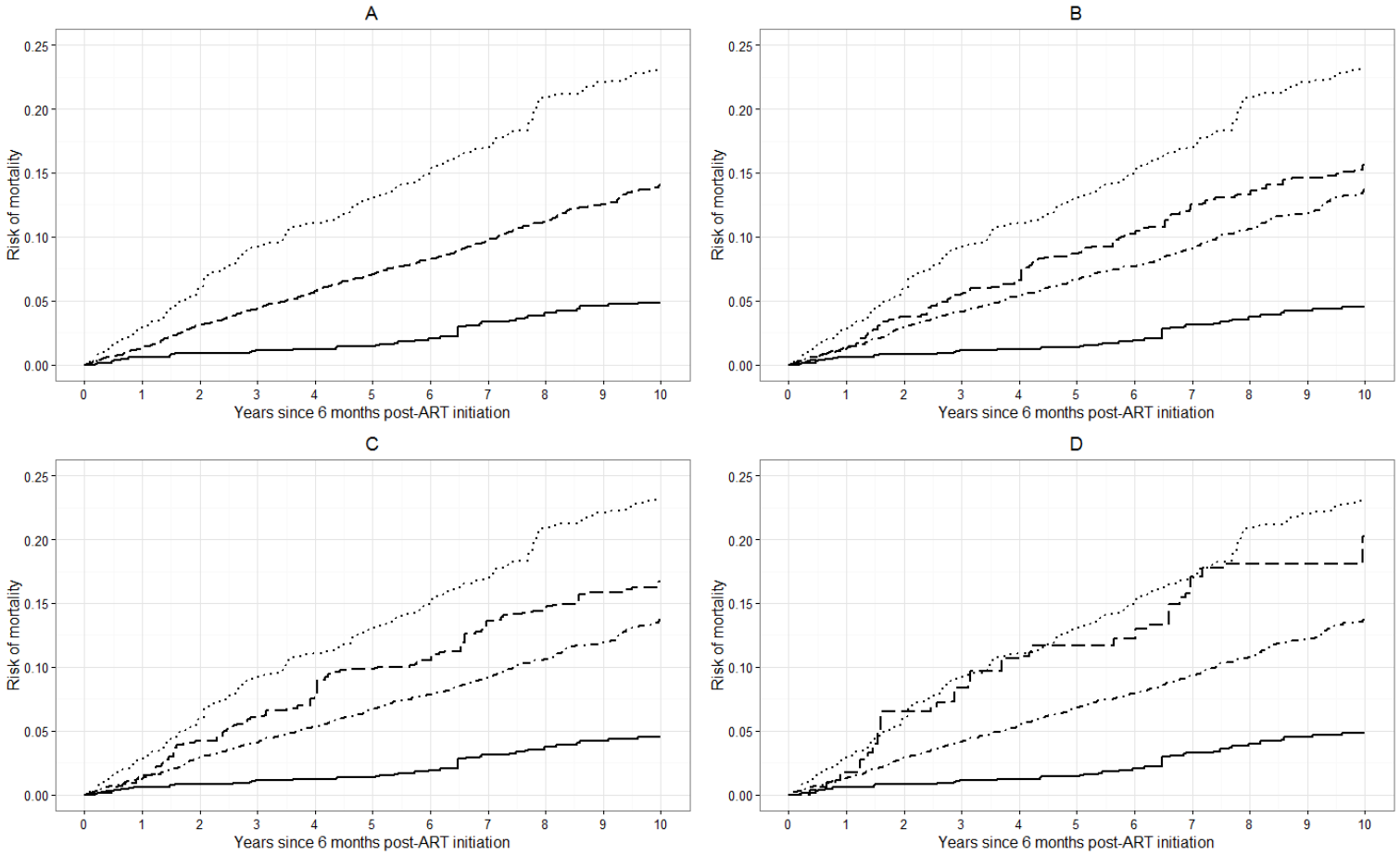
Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site. Upper dashed line indicates hazard ratio for 10-year all-cause mortality for viral loads >999 copies/mL; lower dashed line indicates hazard ratio of 1. A. Crude; B. Standardized.

APPENDIX 4.3D. Crude risk curves for all-cause mortality for 7,944 CNICS patients, for selected viral load threshold values of *k*, stratified by viral load category, with left-censored viral load observations substituted with half of assay detection limit.



Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and *k* copies/mL, dashed line represents viral loads between *k* and 999 copies/mL, dotted line represents viral loads >999 copies/mL. A. *k* = 20 copies/mL; B. *k* = 130 copies/mL; C. *k* = 200 copies/mL; D. *k* = 400 copies/mL.

APPENDIX 4.3E. Standardized risk curves for all-cause mortality for 7,944 CNICS patients, for selected viral load threshold values of *k*, stratified by viral load category, with left-censored viral load observations substituted with half of assay detection limit.



Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, status of other clinical conditions (chronic hepatitis B, chronic hepatitis C, past diagnosis of any cancer excluding nonmelanoma skin cancer), statin use, and study site. Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and *k* copies/mL, dashed line represents viral loads between *k* and 999 copies/mL, dotted line represents viral loads >999 copies/mL. A. *k* = 20 copies/mL; B. *k* = 130 copies/mL; C. *k* = 200 copies/mL; D. *k* = 400 copies/mL.

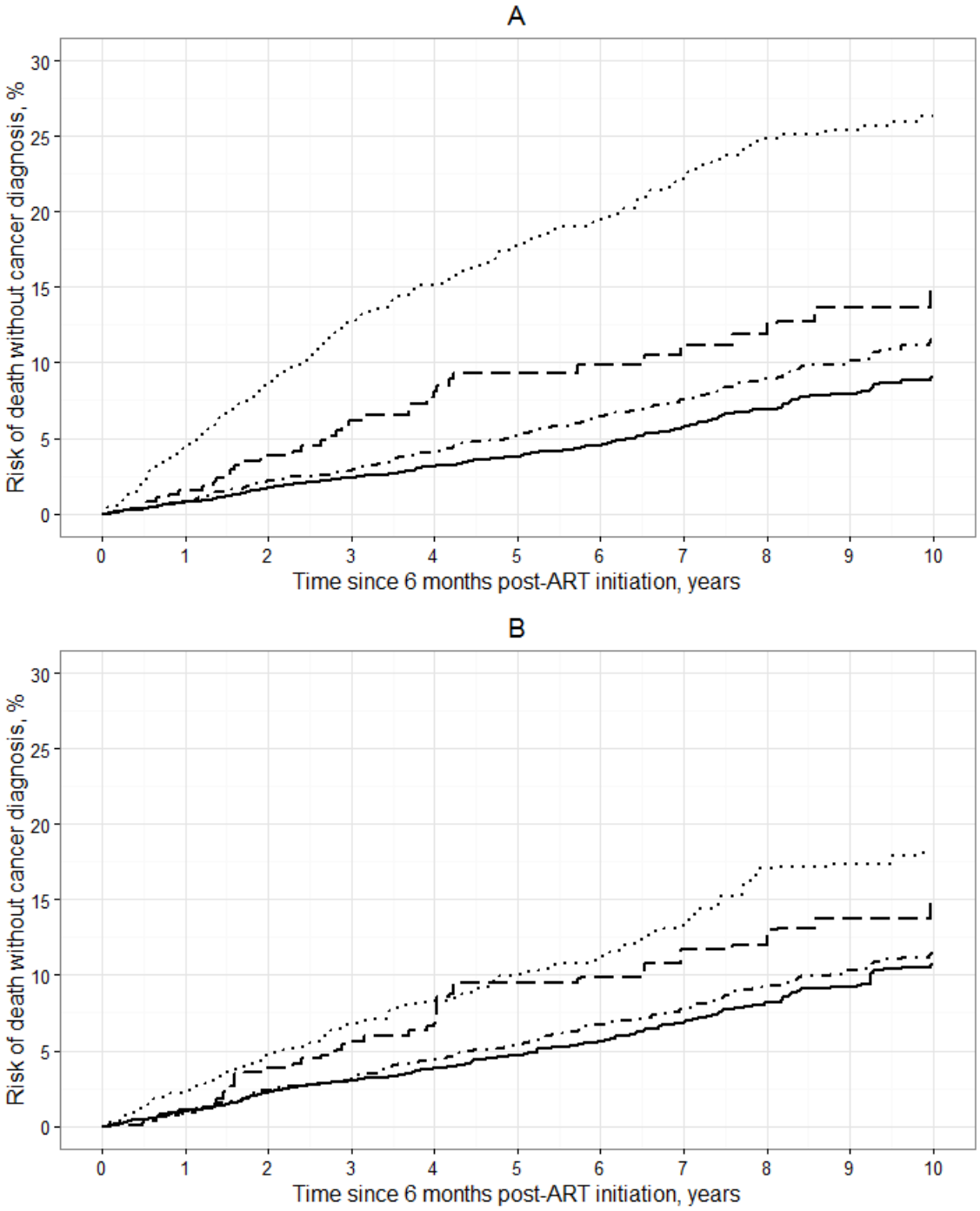
APPENDIX 5.1. Crude and standardized 10-year cumulative incidence, risk difference, and risk ratio estimates for death without a cancer diagnosis in 7,515 CNICS patients, averaged over 30 imputations.

				Crude			Standardized [a]		
	No. of events	No. of patients	Person years	Risk, % (95% CI)	RD, % (95% CI)	RR (95% CI)	Risk, % (95% CI)	RD, % (95% CI)	RR (95% CI)
Total	560	7,515	40,110	12.7 (11.5, 13.9)			12.2 (10.2, 14.2)		
<20 cpm	196	4,281	22,392	8.99 (7.60, 10.4)	0	1	10.7 (8.18, 13.2)	0	1
20 to 199 cpm	112	1,694	9,625	11.5 (9.55, 13.5)	2.55 (0.34, 4.75)	1.28 (1.01, 1.55)	11.5 (9.11, 13.8)	0.74 (-1.68, 3.17)	1.07 (0.83, 1.30)
200 to 999 cpm	37	393	2,124	14.7 (9.55, 19.9)	5.73 (0.29, 11.2)	1.64 (0.99, 2.28)	15.0 (8.86, 21.1)	4.25 (-2.53, 11.0)	1.40 (0.73, 2.06)
>999 cpm	215	1,147	5,969	26.3 (22.9, 29.7)	17.3 (13.6, 21.0)	2.93 (2.31, 3.55)	18.1 (13.7, 22.5)	7.38 (2.40, 12.4)	1.69 (1.14, 2.24)

Abbreviations: CI, confidence interval; cpm, copies per milliliter; RD, risk difference; RR, risk ratio

[a] Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, chronic hepatitis status, statin use, and study site.

APPENDIX 5.2. Crude and standardized risk curves for death without a cancer diagnosis in 7,515 CNICS patients, stratified by viral load category, averaged over 100 imputations.



Standardized estimates controlled for age, sex, race/ethnicity, male-to-male sexual contact, injection drug use, smoking, at-risk alcohol use, pre-ART viral load, year of ART initiation, ART regimen, CD4 count, clinical AIDS status, chronic hepatitis status, statin use, and study site. Solid line represents viral loads <20 copies/mL, dot-dashed line represents viral loads between 20 and 199 copies/mL, dashed line represents viral loads between 200 and 999 copies/mL, dotted line represents viral loads >999 copies/mL. A. Crude; B. Standardized.

REFERENCES

1. Shiels MS, Pfeiffer RM, Gail MH, Hall HI, Li J, Chaturvedi AK, et al. **Cancer burden in the HIV-infected population in the United States**. *Journal of the National Cancer Institute* 2011; 103(9):753-762.

2. Crum-Cianflone N, Hullsiek KH, Marconi V, Weintrob A, Ganesan A, Barthel RV, et al. **Trends in the incidence of cancers among HIV-infected persons and the impact of antiretroviral therapy: a 20-year cohort study**. *AIDS (London, England)* 2009; 23(1):41-50.

3. Sungkanuparph S, Groger RK, Overton ET, Fraser VJ, Powderly WG. **Persistent low-level viraemia and virological failure in HIV-1-infected patients treated with highly active antiretroviral therapy**. *HIV medicine* 2006; 7(7):437-441.

4. Taiwo B, Gallien S, Aga E, Ribaudo H, Haubrich R, Kuritzkes DR, et al. **Antiretroviral drug resistance in HIV-1-infected patients experiencing persistent low-level viremia during first-line therapy**. *The Journal of infectious diseases* 2011; 204(4):515-520.

5. Laprise C, de Pokomandy A, Baril JG, Dufresne S, Trottier H. **Virologic failure following persistent low-level viremia in a cohort of HIV-positive patients: results from 12 years of observation**. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2013; 57(10):1489-1496.

6. Eastburn A, Scherzer R, Zolopa AR, Benson C, Tracy R, Do T, et al. **Association of low level viremia with inflammation and mortality in HIV-infected adults**. *PloS one* 2011; 6(11):e26320.

7. The Antiretroviral Therapy Cohort Collaboration. **Impact of low-level viremia on clinical and virological outcomes in treated HIV-1-infected patients**. *AIDS (London, England)* 2015; 29(3):373-383.

8. Vandenhende MA, Perrier A, Bonnet F, Lazaro E, Cazanave C, Reigadas S, et al. **Risk of virological failure in HIV-1-infected patients experiencing low-level viraemia under active antiretroviral therapy (ANRS C03 cohort study)**. *Antiviral therapy* 2015.

9. Silva J, Pereira K, Rijo J, Alberto T, Cabanas J, Gomes P, et al. **A retrospective observational study of low-level viraemia and its immunological and virological significance: which outcome to expect**. *Journal of the International AIDS Society* 2014; 17(4 Suppl 3):19668.

10. Boillat-Blanco N, Darling KE, Schoni-Affolter F, Vuichard D, Rougemont M, Fulchini R, et al. **Virological outcome and management of persistent low-level viraemia in HIV-1-infected patients: 11 years of the Swiss HIV Cohort Study**. *Antiviral therapy* 2014.

11. Charuratananon S, Sungkanuparph S. **Rate of and predicting factors for virologic failure in HIV-infected patients with persistent low-level viremia under antiretroviral therapy**. *Journal of the International Association of Providers of AIDS Care* 2015; 14(1):12-16.

12. Cohen C. **Low-level viremia in HIV-1 infection: consequences and implications for switching to a new regimen**. *HIV clinical trials* 2009; 10(2):116-124.

13. Delaugerre C, Gallien S, Flandre P, Mathez D, Amarsy R, Ferret S, et al. **Impact of low-level-viremia on HIV-1 drug-resistance evolution among antiretroviral treated-patients**. *PloS one* 2012; 7(5):e36673.

14. World Health Organization. **Global update on the health sector response to HIV, 2014**. Geneva, Switzerland; 2014.

15. World Health Organization. **HIV/AIDS Fact Sheet**. Geneva, Switzerland; 2015.

16. Althoff KN, Buchacz K, Hall HI, Zhang J, Hanna DB, Rebeiro P, et al. **U.S. trends in antiretroviral therapy use, HIV RNA plasma viral loads, and CD4 T-lymphocyte cell counts among HIV-infected persons, 2000 to 2008**. *Annals of internal medicine* 2012; 157(5):325-335.

17. Samji H, Cescon A, Hogg RS, Modur SP, Althoff KN, Buchacz K, et al. **Closing the gap: increases in life expectancy among treated HIV-positive individuals in the United States and Canada**. *PloS one* 2013; 8(12):e81355.

18. Fenton K. **HIV prevention programs among older adult populations**. In: *International AIDS Conference*. Washington, DC; 2012.

19. 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--2012**. *HIV Surveillance Supplemental Report*; 2014.

20. Grulich AE, van Leeuwen MT, Falster MO, Vajdic CM. **Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis**. *Lancet* 2007; 370(9581):59-67.

21. Engels EA. **Non-AIDS-defining malignancies in HIV-infected persons: etiologic puzzles, epidemiologic perils, prevention opportunities**. *AIDS (London, England)* 2009; 23(8):875-885.

22. Robbins HA, Pfeiffer RM, Shiels MS, Li J, Hall HI, Engels EA. **Excess cancers among HIV-infected people in the United States**. *Journal of the National Cancer Institute* 2015; 107(4).

23. Silverberg MJ, Chao C, Leyden WA, Xu L, Tang B, Horberg MA, et al. **HIV infection and the risk of cancers with and without a known infectious cause**. *AIDS (London, England)* 2009; 23(17):2337-2345.

24. Silverberg MJ, Lau B, Achenbach CJ, Jing Y, Althoff KN, D'Souza G, et al. **Cumulative Incidence of Cancer Among Persons With HIV in North America: A Cohort Study**. *Annals of internal medicine* 2015; 163(7):507-518.

25. Chene G, Sterne JA, May M, Costagliola D, Ledergerber B, Phillips AN, et al. **Prognostic importance of initial response in HIV-1 infected patients starting potent antiretroviral therapy: analysis of prospective studies**. *Lancet* 2003; 362(9385):679-686.

26. Achenbach CJ, Buchanan AL, Cole SR, Hou L, Mugavero MJ, Crane HM, et al. **HIV viremia and incidence of non-Hodgkin lymphoma in patients successfully treated with antiretroviral therapy**. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2014; 58(11):1599-1606.

27. Reus S, Portilla J, Sanchez-Paya J, Giner L, Frances R, Such J, et al. **Low-level HIV viremia is associated with microbial translocation and inflammation**. *Journal of acquired immune deficiency syndromes (1999)* 2013; 62(2):129-134.

28. Charpentier C, Landman R, Laouenan C, Joly V, Hamet G, Damond F, et al. **Persistent low-level HIV-1 RNA between 20 and 50 copies/mL in antiretroviral-treated patients: associated factors and virological outcome**. *The Journal of antimicrobial chemotherapy* 2012; 67(9):2231-2235.

29. Do T, Duncan J, Butcher A, Liegler T. **Comparative frequencies of HIV low-level viremia between real-time viral load assays at clinically relevant thresholds**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2011; 52 Suppl 1:S83-89.

30. Cheng CY, Luo YZ, Wu PY, Liu WC, Yang SP, Zhang JY, et al. **Antiretroviral therapy (ART) management of Low-Level Viremia in Taiwan (ALLEVIATE)**. *Journal of the International AIDS Society* 2014; 17(4 Suppl 3):19785.

31. Maggiolo F, Callegaro A, Cologni G, Bernardini C, Velenti D, Gregis G, et al. **Ultrasensitive assessment of residual low-level HIV viremia in HAART-treated patients and risk of virological failure**. *Journal of acquired immune deficiency syndromes (1999)* 2012; 60(5):473-482.

32. Sahu GK. **Potential implication of residual viremia in patients on effective antiretroviral therapy**. *AIDS research and human retroviruses* 2015; 31(1):25-35.

33. Wang S, Rong L. **Stochastic population switch may explain the latent reservoir stability and intermittent viral blips in HIV patients on suppressive therapy**. *Journal of theoretical biology* 2014; 360:137-148.

34. Kitahata MM, Rodriguez B, Haubrich R, Boswell S, Mathews WC, Lederman MM, et al. **Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems**. *International journal of epidemiology* 2008; 37(5):948-955.

35. Panel of Antiretroviral Guidelines for Adults and Adolescents. **Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents**. Department of Health and Human Services; 2016. pp. C5.

36. Armbruster DA, Pry T. **Limit of blank, limit of detection and limit of quantitation**. *The Clinical biochemist Reviews / Australian Association of Clinical Biochemists* 2008; 29 Suppl 1:S49-52.

37. Browne RW, Whitcomb BW. **Procedures for determination of detection limits: application to high-performance liquid chromatography analysis of fat-soluble vitamins in human serum**. *Epidemiology (Cambridge, Mass)* 2010; 21 Suppl 4:S4-9.

38. **1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults**. *MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control* 1992; 41(Rr-17):1-19.

39. Howe CJ, Cole SR, Westreich DJ, Greenland S, Napravnik S, Eron JJ, Jr. **Splines for trend analysis and continuous confounder control**. *Epidemiology (Cambridge, Mass)* 2011; 22(6):874-875.

40. Nie L, Chu H, Liu C, Cole SR, Vexler A, Schisterman EF. **Linear regression with an independent variable subject to a detection limit**. *Epidemiology (Cambridge, Mass)* 2010; 21 Suppl 4:S17-24.

41. Jin Y, Hein MJ, Deddens JA, Hines CJ. **Analysis of lognormally distributed exposure data with repeated measures and values below the limit of detection using SAS**. *The Annals of occupational hygiene* 2011; 55(1):97-112.

42. Gillespie BW, Chen Q, Reichert H, Franzblau A, Hedgeman E, Lepkowski J, et al. **Estimating population distributions when some data are below a limit of detection by using a reverse Kaplan-Meier estimator**. *Epidemiology (Cambridge, Mass)* 2010; 21 Suppl 4:S64-70.

43. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. **Using the outcome for imputation of missing predictor values was preferred**. *Journal of clinical epidemiology* 2006; 59(10):1092-1101.

44. Peto R. **Experimental Survival Curves for Interval-Censored Data**. *Journal of the Royal Statistical Society Series C* 1973; 22(1):86-91.

45. Turnbull BW. **The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data**. *Journal of the Royal Statistical Society, Series B* 1976; 38(3):290-295.

46. Cox DR. **Regression Models and Life-Tables**. *Journal of the Royal Statistical Society, Series B* 1972; 34:187-220.

47. Efron B. **The efficiency of Cox's likelihood function for censored data**. *Journal of the American Statistical Association* 1977; 72:557-565.

48. Rubin DB. **Multiple Imputation for Nonresponse in Surveys**. New York: John Wiley & Sons, Inc.; 1987.

49. Nelson W. **Theory and applications of hazard plotting for censored failure data**. *Technometrics* 1972; 14(4):945-966.

50. Aalen O. **Nonparametric inference for a family of counting processes**. *The Annals of Statistics* 1978; 6(4):701-726.

51. Cole SR, Hernan MA. **Adjusted survival curves with inverse probability weights**. *Computer methods and programs in biomedicine* 2004; 75(1):45-49.

52. Cole SR, Hernan MA. **Constructing inverse probability weights for marginal structural models**. *American journal of epidemiology* 2008; 168(6):656-664.

53. Buchanan AL, Hudgens MG, Cole SR, Lau B, Adimora AA. **Worth the weight: using inverse probability weighted Cox models in AIDS research**. *AIDS research and human retroviruses* 2014; 30(12):1170-1177.

54. Schomaker M, Heumann C. **Bootstrap inference when using multiple imputation**. *arXiv* 2016.

55. Fine JP, Gray RJ. **A proportional hazards model for the subdistribution of a competing risk**. *Journal of the American Statistical Association* 1999; 94(446):496-509.

56. Steigbigel RT, Cooper DA, Kumar PN, Eron JE, Schechter M, Markowitz M, et al. **Raltegravir with optimized background therapy for resistant HIV-1 infection**. *The New England journal of medicine* 2008; 359(4):339-354.

57. Deeks SG. **Determinants of virological response to antiretroviral therapy: implications for long-term strategies**. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2000; 30 Suppl 2:S177-184.

58. Helsel D. **Much ado about next to nothing: incorporating nondetects in science**. *The Annals of occupational hygiene* 2010; 54(3):257-262.

59. Cole SR, Chu H, Nie L, Schisterman EF. **Estimating the odds ratio when exposure has a limit of detection**. *International journal of epidemiology* 2009; 38(6):1674-1680.

60. Xu J, Kochanek KD, Murphy SL, Arias E. **Mortality in the United States, 2012**. In: *NCHS Data Brief*: Centers for Disease Control and Prevention; 2014.

61. Bruyand M, Thiebaut R, Lawson-Ayayi S, Joly P, Sasco AJ, Mercie P, et al. **Role of uncontrolled HIV RNA level and immunodeficiency in the occurrence of malignancy in HIV-infected patients during the combination antiretroviral therapy era: Agence Nationale de Recherche sur le Sida (ANRS) CO3 Aquitaine Cohort**. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2009; 49(7):1109-1116.

62. Silverberg MJ, Chao C, Leyden WA, Xu L, Horberg MA, Klein D, et al. **HIV infection, immunodeficiency, viral replication, and the risk of cancer**. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2011; 20(12):2551-2559.

63. Coussens LM, Werb Z. **Inflammation and cancer**. *Nature* 2002; 420(6917):860-867.

64. Achenbach CJ, Cole SR, Kitahata MM, Casper C, Willig JH, Mugavero MJ, et al. **Mortality after cancer diagnosis in HIV-infected individuals treated with antiretroviral therapy**. *AIDS (London, England)* 2011; 25(5):691-700.

65. Lee JS, Cole SR, Richardson DB, Dittmer DP, Miller WC, Moore RD, et al. **Incomplete viral suppression and mortality in HIV patients after antiretroviral therapy initiation**. *AIDS (London, England)* 2017.

66. Engels EA, Biggar RJ, Hall HI, Cross H, Crutchfield A, Finch JL, et al. **Cancer risk in people infected with human immunodeficiency virus in the United States**. *International journal of cancer* 2008; 123(1):187-194.

67. Gopal S, Patel MR, Achenbach CJ, Yanik EL, Cole SR, Napravnik S, et al. **Lymphoma immune reconstitution inflammatory syndrome in the center for AIDS research network of integrated clinical systems cohort**. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2014; 59(2):279-286.