

## RESEARCH ARTICLE

# Estimation of time of HIV seroconversion using a modified CD4 depletion model

Viviane D. Lima<sup>1,2\*</sup>, Lu Wang<sup>1</sup>, Paul Sereda<sup>1</sup>, Taylor McLinden<sup>1</sup>, Rolando Barrios<sup>1,3</sup>, Julio S. G. Montaner<sup>1,2</sup>

**1** British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada, **2** Division of Infectious Diseases, Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, Canada, **3** Vancouver Coastal Health, Vancouver, Canada

\* [vlima@bccfe.ca](mailto:vlima@bccfe.ca)



## Abstract

### Introduction

Several methods have been proposed to estimate the time of HIV seroconversion, including those based on CD4 cell depletion models. However, previous models have failed to consider the heterogeneity that exists in CD4 trajectories among different sub-populations. Our objective was to estimate the time from HIV seroconversion relative to the HIV diagnosis date in a population-based cohort of people living with HIV (PLWH) in the province of British Columbia, Canada.

### Methods

We used linked administrative and clinical data from the British Columbia Seek and Treat for Optimal Prevention of HIV/AIDS (STOP HIV/AIDS) cohort, which contains longitudinal individual-level data on all PLWH ever diagnosed in the province. Eligible participants were aged  $\geq 18$  years and diagnosed with HIV between 1989 and 2013. The outcome was pre-antiretroviral treatment CD4 cell count measurements assessed every six months. Models were stratified by age and stage of HIV infection at diagnosis. Several explanatory variables were considered including longitudinal viral load measurements. Longitudinal CD4, square root transformed, was modeled via a non-linear mixed effects model; time was modeled using an exponential decay function. We assumed a Gaussian distribution (identity link), an AR(1) correlation structure, and a random intercept and slope for the longitudinal viral load measurements. Due to the population variation in CD4 count among uninfected individuals, we assumed 500 to 1500 cells/mm<sup>3</sup> as the normal range when estimating the time of HIV seroconversion.

### Results

Longitudinal data on 1,253 individuals were analysed: 80% male, 33% White, and the median age at diagnosis was 38 years (25<sup>th</sup>-75<sup>th</sup> percentile [Q<sub>1</sub>-Q<sub>3</sub>], 31 to 45). CD4 decay differed by stage of infection at diagnosis and age, with those  $\geq 50$  years in Stages 1 and 2 experiencing a faster decline in CD4 over time. The median duration of infection from seroconversion until HIV diagnosis was 6.9 (Q<sub>1</sub>-Q<sub>3</sub>, 3.9 to 10.1) years.

## OPEN ACCESS

**Citation:** Lima VD, Wang L, Sereda P, McLinden T, Barrios R, Montaner JSG (2021) Estimation of time of HIV seroconversion using a modified CD4 depletion model. PLoS ONE 16(2): e0246135. <https://doi.org/10.1371/journal.pone.0246135>

**Editor:** Ethan Morgan, Ohio State University, UNITED STATES

**Received:** July 17, 2020

**Accepted:** January 13, 2021

**Published:** February 12, 2021

**Copyright:** © 2021 Lima et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The British Columbia Centre for Excellence in HIV/AIDS is prohibited from making individual-level data available publicly due to provisions in our service contracts, institutional policy, and ethical requirements. The British Columbia Centre for Excellence in HIV/AIDS data is not available externally due to prohibitions in service contracts with our funders or data providers. Institutional policies stipulate that all external data requests require collaboration with a BC Centre for Excellence in HIV/AIDS. For more information regarding data access, please contact Mr. Mark Helberg, the Senior Director, Internal and

External Relations & Strategic Development of the British Columbia Centre for Excellence in HIV/AIDS, at [mhelberg@bccfe.ca](mailto:mhelberg@bccfe.ca). In his position, he oversees some of the daily operations from the Center and any data access request that may arise.

**Funding:** The programmatic funding for the Drug Treatment Program, Treatment as Prevention and STOP HIV/AIDS initiatives was provided by the British Columbia government. Additional research funding was provided by the Canadian Institutes of Health Research (VDL), the Michael Smith Foundation for Health Research (VDL), the Public Health Agency of Canada (JSGM), Vancouver Coastal Health Authority (JSGM), and the Canadian Trials Network (JSGM). The sponsors had no role in the design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data in the study and had final responsibility to submit for publication.

**Competing interests:** Institutional grants have been provided to JSGM by Gilead, J&J, Merck, ViiV Healthcare. The remaining authors do not have conflicts to declare. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Conclusions

Considering the heterogeneity that exists in individual CD4 cell trajectories in a population, we presented a methodology that only relies on routinely collected HIV-related data, which can be further extended to estimate other epidemic measures.

## Introduction

In Canada, as in other high-resource countries, people living with HIV (PLWH) are living longer than ever before due to the success of antiretroviral treatment (ART) [1, 2]. In addition to preventing morbidity and mortality due to HIV, ART has also been shown to stop HIV transmission [3–5]. In order to assess whether countries are moving towards HIV epidemic control [6], it is important to find reliable methods that can be used to estimate important epidemic measures of morbidity, including the time that it takes for an individual infected with HIV to be diagnosed. This information is key to enhancing HIV testing programs and linkage to care, and, therefore, to decreasing health disparities across population subgroups.

The method that we propose in this study is based on CD4 cell count (hereafter referred to CD4) depletion [7–11]. This methodology is flexible enough to consider inherent heterogeneities that exist in a population and it can be further applied to estimate other epidemic measures such as HIV incidence and prevalence. This model is based on longitudinal individual-level information on biomarkers (e.g., CD4, HIV viral load) and demographic factors (e.g., age, sex), and it can be extended to include information on determinants of health (e.g., biological, behaviour and environmental factors) and other factors that are known to influence the natural history of HIV [11]. Therefore, we propose to estimate the duration of infection from HIV seroconversion until diagnosis, age at seroconversion, and year of seroconversion in a population-based cohort of PLWH in the province of British Columbia (BC), Canada using a CD4 depletion model while considering different demographic, clinical and behavioural variables.

## Materials and methods

### Data source and study population

In BC, through the provincial Seek and Treat for Optimal Prevention of HIV/AIDS (STOP HIV/AIDS) population-based retrospective cohort, we have access to longitudinal individual-level data on all PLWH since their date of HIV diagnosis [12, 13]. The STOP HIV/AIDS cohort is based on a data linkage from the BC Centre for Excellence in HIV/AIDS Drug Treatment Program (DTP) clinical registry and several administrative databases containing health information on all diagnosed PLWH (regardless of whether they are accessing ART in BC or not) [13–18]. Since 1992, BC residents living with HIV have had access to centralized and publicly funded ART (through the DTP) and specialized HIV laboratory monitoring, in accordance with the BC Centre for Excellence in HIV/AIDS HIV therapeutic guidelines [19]. Data captured in the STOP HIV/AIDS cohort includes socio-demographic (e.g., sex, age, ethnicity, geographic location of residence), clinical (e.g., CD4, plasma HIV viral load, AIDS-defining illness, mortality), healthcare utilization (e.g., hospitalization, non-ART prescriptions, physician visits) and treatment variables (e.g., antiretroviral regimen information, date of ART initiation). The databases included in the STOP HIV/AIDS cohort, along with their corresponding data capture are comprehensively detailed in the Supplement.

In our study, eligible individuals were aged  $\geq 18$  years at HIV diagnosis, which happened between 1989 and 2013, and they were followed until they started ART treatment between 1996 and 2015, the last contact date with the provincial healthcare system (e.g., a physician visit, hospitalization, laboratory test), the date of death, study end, or the date in which they moved out of BC. Additionally, individuals were required to have at least two measurements of CD4 and viral load during follow-up. All viral load measurements in BC are centrally done at the St. Paul's Hospital virology laboratory. Since the quantification range of viral load assays has evolved over time, for analytical purposes, we truncated our measurements to range from  $< 500$  (coded as 499) to  $> 100,000$  (coded as 100,010) copies/mL [20–23]. CD4 is measured by flow cytometry, followed by fluorescent monoclonal antibody analysis (Beckman Coulter, Inc., Mississauga, Ontario, Canada). The CD4 data are measured at different laboratories across BC; however, we capture  $> 85\%$  of all CD4 tests done in BC in our database. In addition, we removed CD4 values that were outside the normal range for this biomarker (i.e.,  $> 1500$  cells/mm<sup>3</sup>) [24].

### Statistical analysis

CD4 and viral load measurements were obtained every six months from the time since HIV diagnosis until ART initiation, in order to model the CD4 depletion trajectory in our study. Longitudinal CD4, square root transformed, was modeled via a non-linear mixed effects model; follow-up time was modeled using an exponential decay function [25], assuming a Gaussian distribution (identity link), an AR(1) correlation structure, and a random intercept and slope for the longitudinal viral load measurements:

$$\sqrt{CD4_{i,t}} \sim \gamma e^{-R \cdot t_i} + \alpha_{0,i} + \alpha_{1,i} \log_{10}(\text{Viral Load})_{i,t} + \beta_1 x_{1,i} + \dots + \beta_n x_{n,i} + \varepsilon_{i,t}$$

where  $t$  represents each of the 6-month intervals;  $i$  represents each individual in the study;  $\varepsilon_i$  is the random error distributed as  $N(\mathbf{0}, \mathbf{K}_i)$ , where  $\mathbf{K}$  is the covariance matrix and independent of  $\alpha_{0,i}$  and  $\alpha_{1,i}$ , which are the random intercept and slope that varies across individuals, and they follow a bivariable normal distribution  $N\left[(\alpha_0, \alpha_1), (\sigma_{\alpha_0}, \sigma_{\alpha_1}), \mathbf{P}_{\sigma_{\alpha_0}, \sigma_{\alpha_1}}\right]$ , where  $\mathbf{P}$  is the covariance matrix. We also assumed an AR(1) correlation structure for the error term, and an unstructured variance-covariance matrix for the bivariable normal distribution. The coefficients  $\beta_1, \dots, \beta_n$  are for the fixed explanatory variables  $x_1, \dots, x_2$ . The exponential decay function was modeled via the term  $\gamma e^{-R \cdot t}$ , where  $\gamma$  and  $R$  are coefficients in this function. If the exponential decay function was not justified based on model selection, we fitted instead a linear mixed effects model. Model selection was based on a published method by our group, based on Akaike Information Criterion and significance level [26, 27]. Goodness-of-fit was based on residual diagnostic plots (Supplement). Analyses were performed in R© version 3.6.3 using the libraries nlme, mgcv, and ggplot2.

As the CD4 trajectories are expected to be different across population subgroups, the models were stratified by age ( $< 50$  versus  $\geq 50$  years) and stage of HIV infection at diagnosis (Stage 1:  $CD4 \geq 500$ , Stage 2a:  $CD4$  350–499, Stages 2b&3:  $CD4 < 350$  cells/mm<sup>3</sup>) [11]. In the HIV field, we usually define older PLWH as those  $\geq 50$  years [28–30]. In addition, studies have shown that age is an important factor when estimating HIV disease progression using CD4 depletion models [11, 31]. Thus, based on the heterogeneity in CD4 trajectories associated with age, we decided to conduct an age-stratified analysis ( $< 50$  versus  $\geq 50$  years). Note that we did not model individuals whose laboratory criteria indicated acute or recent HIV infection [32]. This decision was made “a priori” based on the established fact that the CD4 during acute HIV infection experience a sharp temporary decline, and estimating time of infection using a methodology based on CD4 would not yield valid results [11, 33]. Explanatory variables

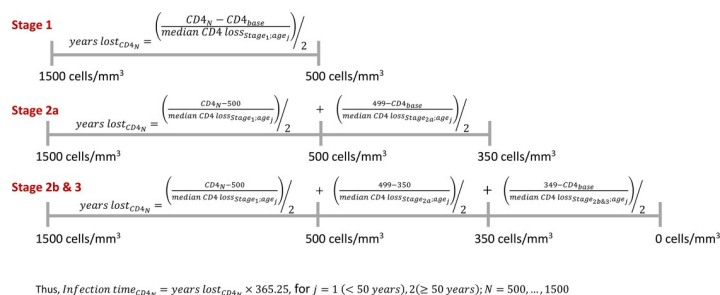
included longitudinal viral load measurements (in log<sub>10</sub>), sex (female/male), year of HIV diagnosis (<1996, 1996–1999, 2000–2003, 2004–2007, 2008–2013), ethnicity (White, non-White, unknown), HIV transmission risk group (gay, bisexual and other men who have sex with men [gbMSM], people who have ever injected drugs [PWID], GBMSM/PWID, heterosexual, other/unknown), AIDS at HIV diagnosis, and follow-up time (years) from HIV diagnosis until ART initiation. Note that we do not have individual-level information on HIV subtype in the STOP HIV/AIDS cohort, but a recent publication using the DTP data estimated that 86% of PLWH in BC have HIV subtype B [34].

The main outcomes from this analysis were the duration of infection from HIV seroconversion until diagnosis as well as age at and year of seroconversion. These outcomes were estimated for each individual based on the statistical model for each age group and stage of HIV infection (six models in total) as follows:

- i. For each model, we estimated the CD4 cell loss per six months from HIV diagnosis until the start of ART. We calculated summary statistics, stratified by key variables, including medians and quartiles.
- ii. Due to the population variation in CD4 cell count among HIV-negative individuals, we assumed 500 to 1500 cells/mm<sup>3</sup> as the normal range [24, 35]. Thus, based on the CD4 value at diagnosis and for each CD4 value in the normal range, we calculated the duration of infection from HIV seroconversion until diagnosis as illustrated in Fig 1.
- iii. Estimates for the year of and age at HIV seroconversion below the lower boundary (*i.*1980, which we assumed to be the year of the first possible infection case in Canada [36, 37]; or *ii.* minimum age of infection of 16 years [38, 39], as very few infections are acquired perinatally in BC and this is the minimum age of sexual consent to sexual activity in Canada) were discarded.
- iv. Given that we had a different estimate for each of our outcomes for each value in the range 500 to 1500 cells/mm<sup>3</sup>, the final results were summarized by medians and quartiles.

### Ethics approval

This study was approved by the University of British Columbia ethics review committee at the St. Paul’s Hospital, Providence Health Care site (H18-02208 and H05-50123). As per the Research Ethics Board approval for this study and in compliance with relevant local legislation, informed consent was not required for this analysis, which was an approved secondary use of the administrative data involved.



**Fig 1. Formula to calculate the number of years lost from HIV sero-conversion and HIV diagnosis, and the time of infection, stratified by stage of HIV infection and age at diagnosis, assuming 500 to 1500 cells/mm<sup>3</sup> as the normal range among uninfected individuals.**

<https://doi.org/10.1371/journal.pone.0246135.g001>

## Results

Longitudinal data for a total of 4591 individuals who received ART for the first time in BC were considered. Of those, 403 (8.8%) individuals were removed as they were considered to be experiencing acute or recent HIV infection, 1066 (23.2%) for having no information on the stage of HIV infection at diagnosis, and 1869 (40.7%) were removed for not meeting the other eligibility criteria (1751 [38.1%] for having only one value for CD4, 117 [2.5%] for having missing information on CD4 and viral load, and one [0.02%] for having a CD4 out of range) (eTable 1 in [S1 File](#)). eTable 2 in [S1 File](#) presents the comparison between individuals included versus excluded from the study. Those included were more likely to be White, gbMSM, diagnosed between 2000 and 2003, with a higher CD4 cell count, without an AIDS diagnosis and a slightly lower viral load.

Thus, in total, 1253 (27.3%) individuals were included in this analysis with a total of 6542 CD4 measurements. The median number of CD4 measurements per individual, between HIV diagnosis and start of ART, was 5.0 (25<sup>th</sup>-75<sup>th</sup> percentile [Q<sub>1</sub>-Q<sub>3</sub>], 3.0 to 9.0) ([Table 1](#)). Of these, 80.0% were male, 33.0% were White, 36.5% gbMSM, 39.0% PWID, 97.0% had no AIDS defining illness at HIV diagnosis and 59.6% were diagnosed before 2004. Additionally, at the time of HIV diagnosis, the median age was 38 (Q<sub>1</sub>-Q<sub>3</sub>, 31 to 45) years, CD4 cell count was 450 (Q<sub>1</sub>-Q<sub>3</sub>, 320 to 610) cells/mm<sup>3</sup>, viral load was 4.6 (Q<sub>1</sub>-Q<sub>3</sub>, 4.0 to 5.0), and these individuals were followed for a median of 2.1 (Q<sub>1</sub>-Q<sub>3</sub>, 1.1 to 4.0) years until ART initiation.

The distribution of individuals in our strata was: 462 (36.9%) in Stage 1 and age <50 years; 64 (5.1%) in Stage 1 and age ≥50 years; 302 (24.2%) in Stage 2a and age <50 years; 51 (4.1%) in Stage 2a and age ≥50 years; 319 (25.5%) in Stages 2b&3 and age <50 years; and 54 (4.3%) in Stages 2b&3 and age ≥50 years. The CD4 cell trajectories for each stage of HIV infection and age stratum are shown in eFig1 in [S1 File](#). The median number of CD4 measurements per individual, for each stratum, was 7.0 (Q<sub>1</sub>-Q<sub>3</sub>, 3.0 to 11.0), 5.0 (Q<sub>1</sub>-Q<sub>3</sub>, 3.0 to 7.0), 5.0 (Q<sub>1</sub>-Q<sub>3</sub>, 3.0 to 9.0), 4.0 (Q<sub>1</sub>-Q<sub>3</sub>, 3.0 to 7.0), 3.0 (Q<sub>1</sub>-Q<sub>3</sub>, 2.0 to 5.0) and 2.0 (Q<sub>1</sub>-Q<sub>3</sub>, 2.0 to 4.0), respectively.

The fitted models and goodness-of-fit assessments can be found in the Supplement. The estimated statistics for CD4 cell depletion per six months are shown in [Fig 2](#). We observed that the median CD4 cell loss for each stratum was 44.2 (Q<sub>1</sub>-Q<sub>3</sub>, 30.5 to 56.5), 49.1 (Q<sub>1</sub>-Q<sub>3</sub>, 19.0 to 2.3), 26.1 (Q<sub>1</sub>-Q<sub>3</sub>, 21.5 to 29.6), 33.4 (Q<sub>1</sub>-Q<sub>3</sub>, 21.3 to 46.6), 15.0 (Q<sub>1</sub>-Q<sub>3</sub>, 13.6 to 16.3) and 12.8 (Q<sub>1</sub>-Q<sub>3</sub>, 11.4 to 14.1) cells/mm<sup>3</sup>, respectively. In Stages 1 and 2a, individuals aged ≥50 years lost CD4 cell counts faster than those <50 years, and this association was reversed in the last stages (2b & 3). Please note that the values for mean and associated 95% confidence interval were very similar.

Thus, using the median CD4 depletion in each stratum, applying the formula in [Fig 1](#), and assuming the normal CD4 range of 500 to 1500 cells/mm<sup>3</sup>, the results for duration of infection from HIV seroconversion until diagnosis as well as age at and year of HIV seroconversion are presented in [Fig 3](#). Note that we could not estimate these outcomes for 13 individuals since their estimated values fell below the stated lower boundary. Based on the estimated outcomes for each CD4 value in the normal range, we obtained that the median year of seroconversion was 1995 (Q<sub>1</sub>-Q<sub>3</sub>, 1991 to 2000), age at seroconversion was 31 (Q<sub>1</sub>-Q<sub>3</sub>, 25 to 39) years and duration of infection from seroconversion until HIV diagnosis was 6.9 (Q<sub>1</sub>-Q<sub>3</sub>, 3.9 to 10.1) years. Those in Stages 1, 2a and 2b-3 were diagnosed after a median of 4.6 (Q<sub>1</sub>-Q<sub>3</sub>, 3.8 to 5.2), 6.9 (Q<sub>1</sub>-Q<sub>3</sub>, 6.1 to 7.7), and 11.0 (Q<sub>1</sub>-Q<sub>3</sub>, 9.3 to 13.4) years after the estimated HIV seroconversion date, respectively. In [Fig 4](#), we present the comparison between the distribution of year of HIV diagnosis and estimated year of seroconversion. Compared with the annual new diagnoses, the peak of seroconversions occurred before 1996 versus 2000–2003 for new diagnoses.

**Table 1. Characteristics of eligible individuals at the time of HIV diagnosis.**

Characteristics at Diagnosis	Study Population
<b>Sex, n(%)</b>	
Male	1002 (80.0)
Female	251 (20.0)
<b>Ethnicity, n(%)</b>	
Non-White	281 (22.4)
White	414 (33.0)
Unknown	558 (44.5)
<b>HIV Acquisition Risk Group, n(%)</b>	
gbMSM	457 (36.5)
PWID	489 (39.0)
gbMSM/PWID	103 (8.2)
Heterosexual	182 (14.5)
Other/Unknown	22 (1.8)
<b>AIDS Defining Illness, n(%)</b>	
No	1216 (97.0)
Yes	37 (3.0)
<b>Year of diagnosis, n(%)</b>	
<1996	67 (5.4)
1996–1999	264 (21.1)
2000–2003	415 (33.1)
2004–2007	251 (20.0)
2008–2013	256 (20.4)
<b>Age (years)</b>	
Median (Q <sub>1</sub> -Q <sub>3</sub> )	37.7 (31.2–45.1)
<b>CD4 cell count (cells/mm<sup>3</sup>)</b>	
Median (Q <sub>1</sub> -Q <sub>3</sub> )	450 (320–610)
<b>Viral load (log<sub>10</sub> copies/mL)</b>	
Median (Q <sub>1</sub> -Q <sub>3</sub> )	4.6 (4.0–5.0)
<b>Follow-up time (years)</b>	
Median (Q <sub>1</sub> -Q <sub>3</sub> )	2.1 (1.1–4.0)
<b>Number of 6-month intervals for CD4 measurements</b>	
Median (Q <sub>1</sub> -Q <sub>3</sub> )	5.0 (3.0–9.0)

Abbreviations: gbMSM—gay, bisexual and other men who have sex with men; PWID—people who have ever injected drugs; Q<sub>1</sub>-Q<sub>3</sub> - 25<sup>th</sup>-75<sup>th</sup> percentiles.

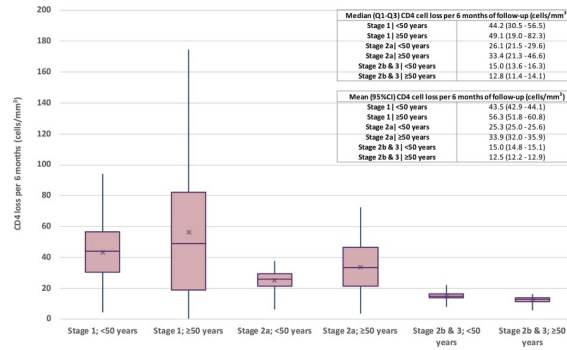
<https://doi.org/10.1371/journal.pone.0246135.t001>

We also observed that the time to HIV diagnosis increased until the 2004–2007 period, and it decreased quite significantly from 2008 to 2013.

## Discussion

In this study, we demonstrated that a model for CD4 cell depletion could be used to estimate the duration of infection from HIV seroconversion until diagnosis, as well as age at and year of seroconversion. We also showed that in order to estimate these parameters, it is important to take into consideration the stage of HIV infection at diagnosis and different demographic, clinical and behavioural variables. The proposed methodology and parameters that we have estimated can be further used to estimate the date of seroconversion of PLWH, in similar settings and demographic profiles as the one in this study, without having to fit this model again.





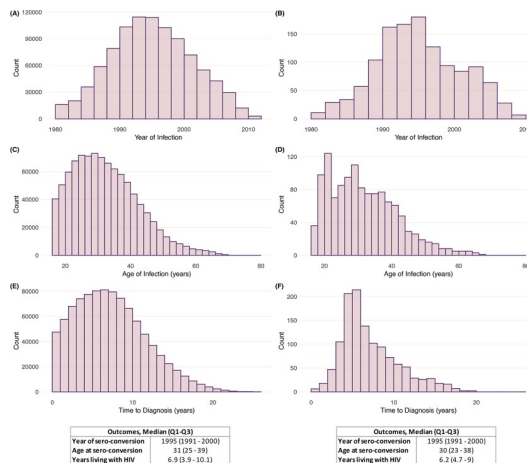
**Fig 2. Estimated CD4 cell loss per six months from HIV diagnosis to start of antiretroviral therapy, stratified by stage of HIV infection and age at diagnosis.**

<https://doi.org/10.1371/journal.pone.0246135.g002>

Another advantage of this method is the possibility of establishing uncertainty around the estimates. However, in order to use this methodology, we stress the importance of obtaining specific information at the time of a HIV diagnosis, especially CD4 cell count and age.

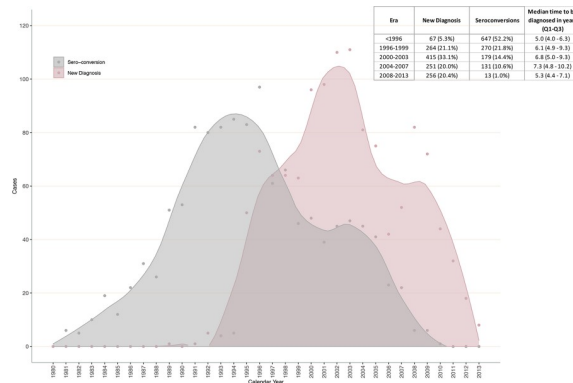
Based on our study, individuals aged  $\geq 50$  years in Stages 1 and 2a ( $CD4 \geq 350$  cells/mm<sup>3</sup>) lost CD4 faster than those  $< 50$  years, and this association was reversed in the Stages 2b & 3 ( $CD4 < 350$  cells/mm<sup>3</sup>). We believe this result was due to the smaller sample size for those aged  $\geq 50$  years in Stages 2b & 3. We also showed that the time to diagnosis varied significantly by year of estimated seroconversion, with a median ranging from 5.0 to 7.3 years before 2008 and 5.3 years thereafter. It is important to remember that during this time, there were key changes in HIV treatment guidelines, shortening the period of time an individual needs to wait between diagnosis and treatment initiation, which could explain these estimates. For example, in 2010 there was the launch of the STOP HIV/AIDS program in 2010, which aims to expand HIV testing, treatment, and support services to all PLWH in BC [2, 40].

As mentioned, our methodology and outputs can be further extrapolated to assist in the estimation of HIV prevalence, which is crucial to monitor the United Nations 90-90-90



**Fig 3. Estimates for the year of seroconversion (A-B), age at seroconversion (C-D), and duration of infection from HIV seroconversion until diagnosis (E-F). Note: For each outcome, the counts on the left are for the estimated values based on each CD4 value in the normal range for each individual (i.e., a maximum of 1000 observations per individual). For each outcome, the counts on the right are for the estimated individual-level values summarized as the median and quartiles of the corresponding distribution on the left (i.e., a maximum of 1 observation per individual).**

<https://doi.org/10.1371/journal.pone.0246135.g003>



**Fig 4. Comparison of the distribution of year of HIV diagnosis and estimated year of HIV seroconversion.**

<https://doi.org/10.1371/journal.pone.0246135.g004>

Targets [41, 42]. These targets propose that, by 2020, at least 90% of PLWH should be diagnosed and aware of their HIV status; at least 90% of those diagnosed be on ART; and at least 90% of those on ART be virologically suppressed. If this target is reached, a 73% virologic suppression coverage will be achieved among all PLWH [41]. Estimating HIV prevalence, which depends on HIV incidence, is not trivial and different methods, depending on data availability, have been proposed. These methods include back-calculations [43–45], next-generation sequencing [46], prevalence surveys [47–49], mathematical modeling [50, 51], and CD4 cell count depletion-based approaches [7–11].

The Public Health Agency of Canada generates biennial national estimates of HIV incidence and prevalence, for the country and for each province and territory, utilizing sophisticated methodologies based on a back-calculation method [43]. Their method relies heavily on routine HIV/AIDS surveillance data (i.e., information on HIV testing, mortality, and recency of HIV infection). Unfortunately, this methodology does not consider demographic, clinical and behavioural heterogeneity that exists among individuals in a population. In addition, this methodology is not very sensitive to changes in policies regarding HIV testing and treatment initiation or the effect of ART in prolonging survival. Consequently, this method can yield estimates with substantial uncertainty, especially in more recent years [52, 53]. Thus, based on a broader population than the one in analyzed in our study, we can use the estimated parameters to reconstruct the HIV epidemic curve, and, ultimately, estimate the HIV prevalence and incidence over time. It would also be possible to estimate the percent of undiagnosed infections for the Canadian population (as well as for provinces and territories) as shown by Song *et al.* for the United States [7].

There are some potential limitations in this study. First, our estimates relied on individuals with at least two measurements of CD4 and viral load. Thus, we did not include individuals with one measurement nor did we use imputation methods to overcome this issue. However, by using the methodology in this study, which considers both intra- and inter-individual variation in the CD4 depletion trajectories (while also adjusting for key explanatory variables and providing goodness-of-fit assessments), we were able to fit robust models for CD4 cell depletion. Second, an important and complicated challenge in this type of analysis is the presence of right truncation of the CD4 data since we stopped following individuals when they initiated treatment. Thus, some individuals would have a shorter follow-up than others, and therefore, fewer CD4 cell count measurements. While mixed effects models largely address this issue, by applying a methodology such as the one by Liu *et al.* or Wu *et al.*, we can further assess the robustness of our findings by examining biases that we may have in our analysis [54, 55].



Third, when estimating our outcomes, we discarded information below the lower boundary established for our estimates. This included assuming a minimum age of infection of 16 years, which excluded 1% of the data not being used to estimate the outcomes. In BC, between 1993 and 2017 there has been less than 40 perinatally acquired HIV infections [38]. In addition, although the minimum age at HIV diagnosis is between 15 and 19 years old, most HIV new diagnoses occurred after the age of 20 years [38]. Thus, although this restriction could have biased our results, we believe that this bias was minimal. Finally, our estimates for CD4 cell depletion considered ethnicity as a variable in the models, despite a large number of individuals in our study being classified as ethnicity unknown. Klein *et al.* showed that individuals of Black ethnicity had a slower rate of CD4 cell depletion than other ethnicities, even after controlling for HIV viral subtype [56]. In our database, only 23 (1.8%) individuals reported having Black ethnicity, and we do not expect that this number will be much higher due to the historical data available during the study period.[38] Although, we could have left out this variable from our model, we let the variable selection method inform whether this variable should be included or not in our final model.

## Conclusions

Considering the heterogeneity that exists in individual CD4 cell trajectories in a population, we presented a methodology that only relies on routinely collected HIV-related data. This methodology yielded robust estimates that can be used in the future to retrospectively estimate other epidemic measures of morbidity, including the proportion of undiagnosed infections, that can be used to assess our progress towards HIV epidemic control in BC.

## Supporting information

**S1 File.**  
(DOCX)

## Acknowledgments

**Disclaimer:** All inferences, opinions, and conclusions drawn in this paper are those of the authors, and do not reflect the opinions or policies of the Data Steward(s).

## Author Contributions

**Conceptualization:** Viviane D. Lima.

**Data curation:** Viviane D. Lima, Lu Wang, Paul Sereda, Taylor McLinden.

**Formal analysis:** Viviane D. Lima, Lu Wang.

**Funding acquisition:** Viviane D. Lima, Rolando Barrios, Julio S. G. Montaner.

**Investigation:** Viviane D. Lima.

**Methodology:** Viviane D. Lima, Lu Wang.

**Project administration:** Viviane D. Lima.

**Resources:** Viviane D. Lima, Rolando Barrios, Julio S. G. Montaner.

**Software:** Viviane D. Lima.

**Visualization:** Viviane D. Lima.

**Writing – original draft:** Viviane D. Lima.

**Writing – review & editing:** Viviane D. Lima, Lu Wang, Paul Sereda, Taylor McLinden, Rolando Barrios, Julio S. G. Montaner.

## References

1. Hogg RS, Eyawo O, Collins AB, Zhang W, Jabbari S, Hull MW, et al. Health-adjusted life expectancy in HIV-positive and HIV-negative men and women in British Columbia, Canada: a population-based observational cohort study. *Lancet HIV*. 2017; 4(6):e270–e6. [https://doi.org/10.1016/S2352-3018\(17\)30029-2](https://doi.org/10.1016/S2352-3018(17)30029-2) PMID: 28262574
2. Lima VD, Brumme ZL, Brumme C, Sereda P, Kraiden M, Wong J, et al. The Impact of Treatment as Prevention on the HIV Epidemic in British Columbia, Canada. *Current HIV/AIDS Rep*. 2020; 17(2):77–87. <https://doi.org/10.1007/s11904-020-00482-6> PMID: 32124189
3. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med*. 2011; 365(6):493–505. <https://doi.org/10.1056/NEJMoa1105243> PMID: 21767103
4. Montaner JS, Hogg R, Wood E, Kerr T, Tyndall M, Levy AR, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. *Lancet*. 2006; 368(9534):531–6. [https://doi.org/10.1016/S0140-6736\(06\)69162-9](https://doi.org/10.1016/S0140-6736(06)69162-9) PMID: 16890841
5. Eisinger RW, Dieffenbach CW, Fauci AS. HIV Viral Load and Transmissibility of HIV Infection: Undetectable Equals Untransmittable. *JAMA*. 2019; 321(5):451–2. <https://doi.org/10.1001/jama.2018.21167> PMID: 30629090
6. Dowdle WR. The principles of disease elimination and eradication. *Bull World Health Organ*. 1998; 76 Suppl 2:22–5. PMID: 10063669
7. Song R, Hall HI, Green TA, Szwarcwald CL, Pantazis N. Using CD4 Data to Estimate HIV Incidence, Prevalence, and Percent of Undiagnosed Infections in the United States. *J Acquir Immune Defic Syndr*. 2017; 74(1):3–9. <https://doi.org/10.1097/QAI.0000000000001151> PMID: 27509244
8. Satcher Johnson A, Song R, Hall HI. Estimated HIV Incidence, Prevalence, and Undiagnosed Infections in US States and Washington, DC, 2010–2014. *J Acquir Immune Defic Syndr*. 2017; 76(2):116–22. <https://doi.org/10.1097/QAI.0000000000001495> PMID: 28700407
9. Farahani M, Novitsky V, Wang R, Bussmann H, Moyo S, Musonda RM, et al. Prognostic Value of HIV-1 RNA on CD4 Trajectories and Disease Progression Among Antiretroviral-Naive HIV-Infected Adults in Botswana: A Joint Modeling Analysis. *AIDS Res Hum Retroviruses*. 2016; 32(6):573–8. <https://doi.org/10.1089/AID.2015.0348> PMID: 26830351
10. Nacher M, Adenis A, Huber F, Hallet E, Abboud P, Mosnier E, et al. Estimation of the duration between HIV seroconversion and HIV diagnosis in different population groups in French Guiana: Strategic information to reduce the proportion of undiagnosed infections. *PLoS One*. 2018; 13(6):e0199267. <https://doi.org/10.1371/journal.pone.0199267> PMID: 29933374
11. Langford SE, Ananworanich J, Cooper DA. Predictors of disease progression in HIV infection: a review. *AIDS Res Ther*. 2007; 4:11. <https://doi.org/10.1186/1742-6405-4-11> PMID: 17502001
12. British Columbia Centre for Excellence in HIV/AIDS (BC-CfE): STOP HIV/AIDS Quarterly Monitoring Report for British Columbia. <http://stophivaids.ca/data-monitoring/> (2020). Accessed 20 March 2020.
13. Heath K, Samji H, Nosyk B, Colley G, Gilbert M, Hogg RS, et al. Cohort profile: Seek and treat for the optimal prevention of HIV/AIDS in British Columbia (STOP HIV/AIDS BC). *International journal of epidemiology*. 2014; 43(4):1073–81. <https://doi.org/10.1093/ije/dyu070> PMID: 24695113
14. British Columbia Centre for Disease Control: Reportable Diseases Data Dashboard. <http://www.bccdc.ca/health-professionals/data-reports/reportable-diseases-data-dashboard> Accessed 8 April 2019.
15. British Columbia Centre for Excellence in HIV/AIDS (BCCfE): Drug Treatment Program. <http://cfenet.ubc.ca/drug-treatment-program> (2019). Accessed 9 October 2019.
16. British Columbia Vital Statistics Agency [creator] (2016):: Vital Statistics Deaths. British Columbia Ministry of Health [publisher]. Data Extract. MOH (2016). <http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central> Accessed.
17. Canadian Institute of Health Information [creator] (2016):: Discharge Abstract Database (Hospital Separations). British Columbia Ministry of Health [publisher]. Data Extract. MOH (2016). <http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central> Accessed.
18. British Columbia Ministry of Health [creator](2016):: Medical Services Plan (MSP) Payment Information File; Consolidation File (MSP Registration & Premium Billing); Home & Community Care (Continuing Care); Mental Health; PharmaNet. British Columbia Ministry of Health [publisher]. Data Extract. MOH

- (2016). <http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central> Accessed.
19. British Columbia Centre for Excellence in HIV/AIDS (BC-CfE): Therapeutic Guidelines for Antiretroviral (ARV) Treatment of Adult HIV Infection. <http://cfenet.ubc.ca/therapeutic-guidelines/guidelines-antiretroviral-arv-treatment-adult-hiv-infection> (2020). Accessed 16 January 2020.
  20. Sun R, Ku J, Jayakar H, Kuo JC, Brambilla D, Herman S, et al. Ultrasensitive reverse transcription-PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 1998; 36(10):2964–9. <https://doi.org/10.1128/JCM.36.10.2964-2969.1998> PMID: 9738051
  21. Schumacher W, Frick E, Kauselmann M, Maier-Hoyle V, van der Vliet R, Babel R. Fully automated quantification of human immunodeficiency virus (HIV) type 1 RNA in human plasma by the COBAS AmpliPrep/COBAS TaqMan system. *J Clin Virol.* 2007; 38(4):304–12. <https://doi.org/10.1016/j.jcv.2006.12.022> PMID: 17329164
  22. Brumme CJ, Swenson LC, Wynhoven B, Yip B, Skinner S, Lima VD, et al. Technical and regulatory shortcomings of the TaqMan version 1 HIV viral load assay. *PLoS one.* 2012; 7(8):e43882. <https://doi.org/10.1371/journal.pone.0043882> PMID: 22937116
  23. De Bel A, Marissens D, Debaisieux L, Liesnard C, Van den Wijngaert S, Lauwers S, et al. Correction of underquantification of human immunodeficiency virus type 1 load with the second version of the Roche Cobas AmpliPrep/Cobas TaqMan assay. *J Clin Microbiol.* 2010; 48(4):1337–42. <https://doi.org/10.1128/JCM.01226-09> PMID: 20164284
  24. HealthLink BC—Healthwise Staff: CD4+ Count. <https://www.healthlinkbc.ca/medical-tests/tu6407#tu6414> (2018). Accessed 23 April 2020.
  25. Lima VD, Le A, Nosyk B, Barrios R, Yip B, Hogg RS, et al. Development and validation of a composite programmatic assessment tool for HIV therapy. *PLoS One.* 2012; 7(11):e47859. <https://doi.org/10.1371/journal.pone.0047859> PMID: 23185244
  26. McCullagh PN, J.A. Generalized Linear Models. 2nd edition ed. Monographs on Statistics and Applied Probability 37. Springer-Science+Business Media, B.V.; 1989.
  27. Pinheiro JCB, D.M. Mixed-Effects Models in S and S-PLUS. Statistics and Computing. New York: Springer-Verlag; 2000.
  28. Hasse B, Ledergerber B, Furrer H, Battegay M, Hirschel B, Cavassini M, et al. Morbidity and aging in HIV-infected persons: the Swiss HIV cohort study. *Clin Infect Dis.* 2011; 53(11):1130–9. <https://doi.org/10.1093/cid/cir626> PMID: 21998280
  29. Jacobson SA. HIV/AIDS interventions in an aging U.S. population. *Health Soc Work.* 2011; 36(2):149–56. <https://doi.org/10.1093/hsw/36.2.149> PMID: 21661304
  30. Womack JA, Justice AC. The OATH Syndemic: opioids and other substances, aging, alcohol, tobacco, and HIV. *Curr Opin HIV AIDS.* 2020; 15(4):218–25. <https://doi.org/10.1097/COH.0000000000000635> PMID: 32487817
  31. Touloumi G, Pantazis N, Pillay D, Paraskevis D, Chaix ML, Bucher HC, et al. Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts. *Clin Infect Dis.* 2013; 56(6):888–97. <https://doi.org/10.1093/cid/cis1000> PMID: 23223594
  32. British Columbia Centre for Disease Control: Communicable Disease Control Manual. Chapter 5—Sexually Transmitted Infections. Guidelines for Testing, Follow up, and Prevention of HIV. [http://www.bccdc.ca/resource-gallery/Documents/Communicable-Disease-Manual/Chapter%205%20-%20STI/HIV\\_Guidelines\\_Testing\\_FollowUp\\_Prevention.pdf](http://www.bccdc.ca/resource-gallery/Documents/Communicable-Disease-Manual/Chapter%205%20-%20STI/HIV_Guidelines_Testing_FollowUp_Prevention.pdf) (2016). Accessed 20 March 2020.
  33. Daar ES, Pilcher CD, Hecht FM. Clinical presentation and diagnosis of primary HIV-1 infection. *Curr Opin HIV AIDS.* 2008; 3(1):10–5. <https://doi.org/10.1097/COH.0b013e3282f2e295> PMID: 19372938
  34. Rocheleau G, Brumme CJ, Shoveller J, Lima VD, Harrigan PR. Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. *Clin Microbiol Infect.* 2018; 24(2):185–91. <https://doi.org/10.1016/j.cmi.2017.06.014> PMID: 28652115
  35. U.S. Department of Veterans Affairs: CD4 count (or T-cell count). <https://www.hiv.va.gov/patient/diagnosis/labs-CD4-count.asp> Accessed 23 April 2020.
  36. CATIE: A history of HIV/AIDS. <https://www.catie.ca/en/world-aids-day/history> (2019). Accessed 21 May 2020.
  37. Canadian AIDS Society: History of Canadian AIDS Society. <https://www.cdnaids.ca/about-us/history/> (2020). Accessed 21 May 2020.
  38. British Columbia Centre for Disease Control: HIV/AIDS Reports. <http://www.bccdc.ca/health-professionals/data-reports/communicable-diseases/hiv-aids> Accessed 28 April 2020.
  39. Government of Canada Department of Justice: Age of Consent to Sexual Activity. <https://www.justice.gc.ca/eng/rp-pr/other-autre/clp/faq.html> (2017). Accessed 21 May 2020.

40. British Columbia Centre for Excellence in HIV/AIDS (BCCfE): Seek and Treat for Optimal Prevention of HIV/AIDS (STOP HIV/AIDS). <http://stophivaids.ca> (2019). Accessed 9 October 2019.
41. The Joint United Nations Programme on HIV and AIDS (UNAIDS): 90-90-90 An ambitious treatment target to help end the AIDS epidemic. [https://www.unaids.org/sites/default/files/media\\_asset/90-90-90\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf) (2014). Accessed 1 December 2018.
42. United Nations: The Millennium Development Goals Report. <https://www.undp.org/content/undp/en/home/librarypage/mdg/the-millennium-development-goals-report-2015.html> (2015). Accessed 17 March 2020.
43. Yan P, Zhang F, Wand H. Using HIV Diagnostic Data to Estimate HIV Incidence: Method and Simulation. *Statistical Communications in Infectious Diseases* 2011.
44. European Centre for Disease Prevention and Control: ECDC HIV Modelling Tool User Manual. <https://www.ecdc.europa.eu/sites/portal/files/documents/HIV-modelling-tool-v1.3.0-manual.pdf> (2017). Accessed 17 March 2020.
45. Wong NS, Wong KH, Lee MP, Tsang OT, Chan DP, Lee SS. Estimation of the Undiagnosed Intervals of HIV-Infected Individuals by a Modified Back-Calculation Method for Reconstructing the Epidemic Curves. *PLoS One*. 2016; 11(7):e0159021. <https://doi.org/10.1371/journal.pone.0159021> PMID: 27403882
46. Puller V, Neher R, Albert J. Estimating time of HIV-1 infection from next-generation sequence diversity. *PLoS Comput Biol*. 2017; 13(10):e1005775. <https://doi.org/10.1371/journal.pcbi.1005775> PMID: 28968389
47. Stover J. Projecting the demographic consequences of adult HIV prevalence trends: the Spectrum Projection Package. *Sex Transm Infect*. 2004; 80 Suppl 1:i14–8. <https://doi.org/10.1136/sti.2004.010157> PMID: 15249694
48. Stover J, Glaubius R, Mofenson L, Dugdale CM, Davies MA, Patten G, et al. Updates to the Spectrum/AIM model for estimating key HIV indicators at national and subnational levels. *AIDS*. 2019; 33 Suppl 3: S227–S34. <https://doi.org/10.1097/QAD.0000000000002357> PMID: 31805028
49. Presanis AM, Gill ON, Chadborn TR, Hill C, Hope V, Logan L, et al. Insights into the rise in HIV infections, 2001 to 2008: a Bayesian synthesis of prevalence evidence. *AIDS*. 2010; 24(18):2849–58. <https://doi.org/10.1097/QAD.0b013e32834021ed> PMID: 20962617
50. Lima VD, Johnston K, Hogg RS, Levy AR, Harrigan PR, Anema A, et al. Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. *J Infect Dis*. 2008; 198(1):59–67. <https://doi.org/10.1086/588673> PMID: 18498241
51. van Sighem A, Nakagawa F, De Angelis D, Quinten C, Bezemer D, de Coul EO, et al. Estimating HIV Incidence, Time to Diagnosis, and the Undiagnosed HIV Epidemic Using Routine Surveillance Data. *Epidemiology*. 2015; 26(5):653–60. <https://doi.org/10.1097/EDE.0000000000000324> PMID: 26214334
52. Brookmeyer R. Measuring the HIV/AIDS epidemic: approaches and challenges. *Epidemiol Rev*. 2010; 32:26–37. <https://doi.org/10.1093/epirev/mxq002> PMID: 20203104
53. Hallett TB. Estimating the HIV incidence rate: recent and future developments. *Curr Opin HIV AIDS*. 2011; 6(2):102–7. <https://doi.org/10.1097/COH.0b013e328343bfdb> PMID: 21505383
54. Liu M, Ying Z. Joint analysis of longitudinal data with informative right censoring. *Biometrics*. 2007; 63(2):363–71. <https://doi.org/10.1111/j.1541-0420.2007.00708.x> PMID: 17425632
55. Yu T, Wu L, Gilbert PB. A joint model for mixed and truncated longitudinal data and survival data, with application to HIV vaccine studies. *Biostatistics*. 2017; 19(3):374–90. <https://doi.org/10.1093/biostatistics/kxx047> PMID: 29028943
56. Klein MB, Young J, Dunn D, Ledergerber B, Sabin C, Cozzi-Lepri A, et al. The effects of HIV-1 subtype and ethnicity on the rate of CD4 cell count decline in patients naive to antiretroviral therapy: a Canadian-European collaborative retrospective cohort study. *CMAJ Open*. 2014; 2(4):E318–29. <https://doi.org/10.9778/cmajo.20140017> PMID: 25485259