

Development and validation of a mortality risk prediction index score for adults living with HIV and multiple chronic comorbidities

Viviane D. Lima^{1,2*}, Bronhilda T. Takeh¹, Neil Fought¹, Hasan Nathani¹, Jieliu Zhu¹, Scott Emerson¹, Katerina Dolguikh¹, Jason Trigg¹, Kate A. Salters¹, Rolando Barrios¹, and Julio S. G. Montaner^{1,2}

Submission History

Submitted:	10/10/2024
Accepted:	04/04/2025
Published:	10/06/2025

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

Abstract

Introduction

Aging while living with HIV poses new challenges in clinical management, mainly due to the onset of multiple chronic comorbidities. Population-specific risk prediction indices considering comorbidities and other risk factors are essential to comprehensively characterise disease burden among PLWH. We developed and validated a mortality risk prediction index (MRP_i) to predict the risk of one-year all-cause mortality among people living with HIV (PLWH).

Methods

Participants were ≥ 18 years and had initiated antiretroviral therapy (ART) between 01/2001 and 12/2018, in British Columbia, Canada. The index date was randomly selected between one-year post-ART initiation and the end of the follow-up. Participants were followed for at least one year from the index date until 12/2019, the last contact date, or the date of death (all-cause), whichever came first. The MRP_i included 18 physical/mental comorbidities, demographic and clinical variables, and ranged from 0 (no risk) to 100 (highest risk).

Results

The final model demonstrated the highest discrimination (c-statistic 0.8355, 95% CI: 0.8187-0.8523 in the training dataset and 0.7965, 95% CI: 0.7664-0.8266 in the test dataset). The comorbidities with the highest weights in the MRP_i were substance use disorders, metastatic solid tumors and non-AIDs defining cancers. For example, for an MRP_i of 30, the predicted one-year all-cause mortality was 0.2%, while an MRP_i of 50 had a predicted mortality of 2.3%.

Conclusions

The MRP_i provides a promising tool to assess the risk of short-term mortality among PLWH in the modern ART era that can inform clinical practice and health policy decisions.

Keywords

aging; comorbidity; mortality; burden of disease; mortality risk prediction; validation

*Corresponding Author:

Email Address: vlima@bccfe.ca (Viviane D. Lima)



Introduction

Following advancements in antiretroviral therapy (ART), there has been a shift in the prognosis of HIV infection from an acute, fatal illness to a manageable, yet complex, chronic condition [1]. As a result, the incidence of AIDS-defining events has declined [2], and the life expectancy of people living with HIV (PLWH) has improved approaching that of the general population [3, 4]. Consequently, many high-resource settings are experiencing a significant demographic shift among PLWH. For example, in British Columbia (BC), Canada, the proportion of diagnosed PLWH ≥ 50 years has increased from 47.0% in 2013 to 65% in 2022 [5, 6].

Aging while living with HIV poses new challenges in clinical management, mainly due to the onset of multiple chronic comorbidities. In BC, the proportion of PLWH with at least one comorbid condition, such as cardiovascular disease (CVD), chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), non-AIDS-defining cancers (NADC), and chronic liver disease (CLD), is higher than their demographically similar HIV-negative counterparts [7]. Additionally, chronic inflammation, uncontrolled viremia, socioeconomic status, and lifestyle factors (e.g., substance use) exacerbate the risk of comorbidities among PLWH [1, 8, 9]. These chronic comorbidities are largely responsible for the continued excess mortality observed in this population, compared with the general population, despite the success of ART [10]. Clinical management of comorbidities may lead to drug-drug interactions and higher healthcare utilisation and costs, which represent an additional challenge for the individual and an added pressure for the healthcare system [11, 12]. Therefore, population-specific risk prediction weighted indices considering comorbidities and other risk factors are essential to comprehensively characterise disease burden among PLWH.

Previous validated weighted indices in the general population include the Charlson Comorbidity Index, Elixhauser Comorbidity Index, and the John Hopkins Adjusted Clinical Group Case-Mix System [13–15]. While these indices consider comorbid conditions in the general adult population, they are not appropriate to use among PLWH. However, there are few indices specifically for PLWH. In the past, risk prediction indices were derived from traditional biomarkers, such as HIV plasma viral load (pVL), CD4, and the presence of AIDS-defining events, all of which had limitations in their ability to accurately reflect the effects of HIV and ART on morbidity and mortality [16, 17]. In the modern ART era, the Veterans Aging Cohort Study (VACS) index 1.0 was derived based on HIV and non-HIV biomarkers and was later updated to the VACS index 2.0 with more variables and a superior ability to predict mortality and other adverse health outcomes [18–20]. Although evidence supports the accuracy of the VACS index in predicting outcomes in PLWH [21], some of its biomarkers (e.g., fibrosis-4, albumin, white blood count) are not routinely monitored by most ART programs and are not available in patient registries [19]. With the increasing use of administrative data in health research [22], a population-specific risk prediction weighted index relevant to this data source would be best for widespread applicability. Therefore, we sought to develop and validate a mortality risk prediction index (MRPi) to predict the risk of one-year all-cause mortality

among PLWH in BC. Although long-term (e.g., five-year) mortality predictions can offer broader insights, there is a critical need for short-term (one-year) risk assessment in clinical practice. Many individuals with multiple comorbidities require urgent interventions, and clinicians often need to prioritise resources and treatment decisions based on the likelihood of near-term adverse outcomes. Therefore, a one-year mortality risk tool is particularly relevant for guiding immediate care strategies and ensuring timely support for those most at risk.

Methods

Data source

The Seek and Treat for Optimal Prevention of HIV/AIDS (STOP HIV/AIDS) cohort is a de-identified population-based cohort of all diagnosed PLWH in BC, followed between April 1st, 1996, and March 31st, 2020. The STOP HIV/AIDS cohort was formed through the annual linkage between the BC Centre for Excellence in HIV/AIDS' Drug Treatment Program registry and other provincial administrative databases, as described in Supplementary Table S1 [23]. It is important to mention that, in BC, ART and related medical and laboratory monitoring are available free of charge to all PLWH residing in the province, which is likely to minimise any bias related to access to care in our analyses.

Study participants

Eligible participants met the following criteria: (i) initiated ART between January 1, 2001, and December 31, 2018, in BC; (ii) were ≥ 18 years old at the first ART date; (iii) had >1 pVL measurement between the first ART date and the end of follow-up; (iv) were ART naïve at baseline; (v) had a pVL ≥ 50 copies/mL at the first ART date; (vi) had >1 year of follow-up after ART initiation; (vii) had ≥ 5 years of administrative data before the randomly selected index date to ascertain comorbidity status [24]; and (viii) had >1 CD4 measurement between 1.5 years before and 6 months after the index date. The index date itself was chosen at least one year after the participant's ART initiation (i.e., between one-year post-ART and the end of follow-up). Index dates were chosen randomly so that results from the final model would not be relative to a specific clinical visit date. We excluded PLWH who had documented structured treatment interruptions or participated in blinded trials during follow-up (see Supplementary Figure S1). These criteria were used to enhance the reproducibility of our study and minimise potential biases.

We followed all eligible participants for >1 year from the index date until i) December 31st, 2019; ii) the last contact date (i.e., the last filled ART prescription refill date, the last available laboratory test date or the date of the last interaction with the healthcare system); or iii) the date of death (all-cause), whichever came first. Baseline pVL and CD4 nadir (lowest CD4) values were obtained within the above-mentioned window period. We omitted anyone missing a CD4 nadir measurement within the window period. Supplementary Figure S1 outlines the step-by-step process for inclusion in the final analytic sample.

Outcome and predictors

The outcome of interest was a one-year all-cause mortality probability. The mortality data was obtained from the BC Vital Statistics Agency database, available through STOP HIV/AIDS [25]. We chose chronic comorbidities: CVD, CLD, COPD, CKD, diabetes mellitus (DM), hypertension (HTN), substance use disorders (SUD), Alzheimer's/dementia (AD/D), NADC, personality disorder (PD), schizophrenia (SCZ), mood/anxiety disorder (MAD), rheumatoid arthritis (RA), metastatic solid tumor (MST), osteoporosis (OP), asthma, peptic ulcer (PU), and osteoarthritis (OA), which are responsible for the excess mortality observed among PLWH despite the success of ART [7, 10]. These comorbidities were assumed to be irreversible once identified and considered present until the end of follow-up. In line with recent recommendations [26], we selected a broad list of 18 comorbidities that are most prevalent and clinically impactful among people living with HIV. We used administrative data definitions for each condition (see Supplementary Table S2), ensuring that we captured a wide spectrum of chronic comorbidities. We acknowledge that other conditions may also be relevant; future work with more comprehensive data could explore additional comorbidities.

The prevalence of each comorbidity was determined using a five-year lookback window from the index date [24], which is consistent with recommendations from Nanditha et al. [24], who demonstrated that longer windows substantially improve the accuracy of identifying chronic diseases in administrative data. Although a shorter window (e.g., two years) might have retained more participants, it risks missing previously diagnosed, but currently stable, conditions. Therefore, we opted for five years to minimise misclassification of comorbidities, aligning with other validated comorbidity indices. We acknowledge that this requirement reduced the eligible sample, but we believe it provides a more robust identification of chronic comorbidities. We ascertained diagnoses of these comorbidities based on the BC Ministry of Health (BC-MoH) case-finding algorithms or literature (where applicable) using the International Classification of Diseases (specifically the ninth revision with clinical modifications, and the Canadian tenth revision version) and Drug/Product Identification Numbers (Supplementary Table S2) [27]. Due to limitations inherent in administrative datasets, we were unable to capture the severity levels or staging of these chronic comorbidities. Consequently, comorbidities in our models were treated as binary variables (present vs. absent).

Other predictors included sex at birth (male, female), age (continuous), years on ART (i.e., years since ART initiation; continuous), pVL suppression (i.e., all pVL measurements <200 copies/mL; yes/no), CD4 nadir (<50, 50-199, 200-349, ≥350 cells/mm³; or continuous), and income assistance (yes/no), measured within the window period mentioned above. In multivariable analysis, we considered linear, quadratic, and cubic polynomial forms for age per 10-year increase, years on ART, and CD4 nadir per 100 cells/mm³, and we centred these variables at the median to control for multicollinearity. Due to a high proportion of missing data for ethnicity in our database, we used 'income assistance' as a proxy for socioeconomic status.

Statistical analyses

Categorical variables were expressed as counts and percentages while continuous variables were expressed as median values with 25th to 75th percentiles (Q1-Q3). In bivariable analyses, we compared categorical variables using the Chi-Square or Fisher's exact test under the applicable conditions and continuous variables using the Wilcoxon rank-sum test [28]. We did not have an independent cohort for external validation. Thus, we performed internal validation using data-splitting (or split-sample) and a bootstrapping-based method and presented both ways. This method accounts for the over-optimism of using unadjusted bootstrapping results and adjusts the model for overfitting [29]. Cox proportional hazard regression models were used to derive the MRP_i to predict one-year all-cause mortality probability. The coefficients of the multivariable models were presented as adjusted hazard ratios (aHRs) with 95% Wald Confidence Intervals (CI). Analyses were performed using SAS version 9.4 (SAS Institute, Inc. Cary, NC, USA) and R statistical software version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria). All p-values were double-sided, and significance was at the 5% level.

Model development

To improve readability and reduce redundancy, we provide a concise overview of the model-building steps here, while detailed derivation procedures are given in Supplementary Text S1.

Data-splitting method

We randomly split the sample into training and test datasets with a ratio 2:1 [29]. The training dataset was used for the model development and later validated on the test dataset. We built Cox proportional hazard regression models on the training dataset to assess the relationship between selected predictor variables and one-year all-cause mortality. We used a variable selection approach based on the Akaike information criterion and Type III p-value [30]. We considered the possibility that certain predictors (e.g., age, CD4 nadir, viral load suppression) might interact with comorbidities to influence one-year all-cause mortality. Accordingly, we tested these interaction terms in preliminary models. However, none of these interactions reached statistical significance or meaningfully improved model discrimination (Harrell's c-statistic) or calibration. To maintain parsimony and interpretability, we, therefore, excluded interaction terms from the final model. We adopted a three-model approach to systematically assess the incremental predictive gain of each additional set of covariates. Model 1 included only comorbidities, Model 2 added age and sex at birth, and Model 3 further incorporated HIV-specific variables (e.g., plasma viral load suppression) and socioeconomic status (income assistance). This stepwise strategy allowed us to evaluate how much discrimination and calibration improved with each new set of predictors.

MRP index score derivation and mortality probability estimation

We used the coefficients derived from the chosen final predictive model (from the data-splitting method) to create

the MRP_i. First, we multiplied each coefficient by the value of the corresponding predictor to get a risk coefficient. Then, we summed the risk coefficients for each participant. Next, participants' MRP_i scores were calculated as shown in Supplementary Text S1. Scores range from 0 to 100 (higher values = higher outcome risk). Next, the predicted one-year all-cause mortality probability was computed using the baseline Cox survival function and the sum of risk coefficients [29]. To compute each participant's one-year all-cause mortality probability, we applied the baseline Cox survival function to the sum of each individual's risk coefficients. The full equation, including the baseline hazard and parameter estimates, is provided in Supplementary Text S1. Also, to facilitate clinical use, we developed both a web-based calculator and an Excel-based tool to estimate the MRP_i and the corresponding one-year all-cause mortality probability. The web-based calculator provides a user-friendly interface, allowing clinicians to enter key patient characteristics to obtain an immediate mortality risk estimate. We have included the link to the web-based calculator and instructions for using the Excel tool in the Supplementary Data File 2 (Mortality Risk Prediction Calculator).

Model performance

We assessed the model's performance through discrimination and calibration. For predictive discrimination, we calculated Harrell's concordance statistic (c-statistic) [31]. The c-statistic ranges from 0.5 to 1.0, with 0.5 considered discrimination by chance alone, 0.70-0.79 considered good, and ≥ 0.80 excellent [32]. We calculated the 95% CI for each c-statistic, assuming a normal distribution [28]. The final predictive model was the model with the best discriminative ability. For calibration, we assessed the agreement between the predicted mortality probability and observed mortality probability using a calibration plot [29].

Data-splitting internal validation

The Cox model coefficients estimated from the training dataset were applied to the test dataset along with the estimation of the c-statistic. Calibration was performed as follows. We first predicted one-year all-cause mortality using our final Cox model derived from the training dataset [20]. Second, we binned the participants in the test dataset into ten subgroups based on a 5-point interval of their predicted mean MRP_i [20]. Next, we estimated the observed one-year all-cause mortality probabilities using the Kaplan-Meier (KM) method and 95% CI for each subgroup. Finally, we graphically compared the predicted and the observed one-year all-cause mortality probabilities for the subgroups. The agreement between the predicted curve and the observed points indicates how well the model was calibrated.

Bootstrapping-based internal validation

The bootstrapping-based internal method estimates the future performance of the model on new participants by providing bias-corrected estimates when the model is applied to a new sample [29]. We carried out this intuitive approach to emphasise the internal validity of the MRP_i. To perform the

bootstrapping-based internal validation, we first built a Cox model using the overall dataset, and the best model was chosen based on its performance, as described above. We estimated c-statistics and predicted survival probabilities by applying the model to the full cohort. Then, we drew 400 bootstrap samples with replacement from the full cohort; for each bootstrap sample, we fitted a Cox model to measure its apparent performance. The optimism was estimated following the method by Harrell *et al.* [29, 31]. We then used the average of the optimism over all 400 bootstrap samples as a correction factor to estimate a validated performance (optimism-adjusted) measure by subtracting the estimated optimism from the apparent performance. For the calibration plot, we calculated and compared the observed and adjusted KM (y-axis) to the predicted (x-axis) one-year survival probabilities, where participants were binned into ten subgroups using deciles of their predicted one-year survival probabilities [29]. A 45-degree line with a slope equal to one and intercept equal to zero (perfect calibration) indicates how well the model was calibrated. Any deviation above or below this line implies the difference between the observed and predicted on year survival probabilities.

Secondary analyses

First, we compared participants included in this study with those excluded to address external validity. Next, we performed sex-based analyses to assess whether the performance of the MRP_i varied by sex at birth (male vs. female). Specifically, we repeated the final model-building steps described above, stratifying by sex.

Results

Population characteristics

Table 1 shows the descriptive characteristics of the full cohort, comparing the training and test datasets at the index date. Among the 4,387 participants, 3,570 (81%) were male, 3,116 (71%) had a suppressed pVL, and 762 (18%) had a CD4 nadir < 200 cells/mm³. The median age was 47 years (Q1-Q3: 38-54), follow-up time was 3.3 years (Q1-Q3: 1.4-6.3), and time on ART was 4.0 years (Q1-Q3: 2.0-6.0). MAD was the most prevalent comorbidity (2,154, 49%), followed by SUD (1,720, 39%), CLD (735, 17%), asthma (596, 14%), HTN (561, 13%), and the least prevalent, RA (38, 1%). Participants in the test dataset and training sets were very comparable, except for asthma prevalence at the index date where those in the test dataset were significantly more likely to have asthma than those in the training dataset (Table 1).

Overall, 714 (16%) deaths were observed with a mortality rate of 38.18 per 1,000 PYs (95% CI: 35.46-41.06 per 1,000 PYs) by the end of follow-up, and 54.09 per 1,000 PYs (95% CI: 47.22-61.70 per 1,000 PYs) in the first year. The KM survival curve from the index date to the end of the study period is shown in Supplementary Figure S2.

Model development and performance

Table 2 summarises the multivariable-adjusted association of the predictors with one-year all-cause mortality and

Table 1: Descriptive characteristics of the full cohort and the cohort participants included in training and test datasets for people living with HIV in British Columbia, Canada, from 2002 to 2019

Variables	Full cohort (n = 4387) n (%)	Test dataset (n = 1463) n (%)	Training dataset (n = 2924) n (%)	p-value
Sex at birth				
Female	817 (19)	284 (19)	533 (18)	0.3637
Male	3570 (81)	1179 (81)	2391 (82)	
Death during the study period				
No	3673 (84)	1238 (85)	2435 (83)	0.2740
Yes	714 (16)	225 (15)	489 (17)	
Cardiovascular disease				
No	4032 (92)	1355 (93)	2677 (92)	0.2456
Yes	355 (8)	108 (7)	247 (8)	
Chronic kidney disease				
No	4095 (93)	1362 (93)	2733 (93)	0.6883
Yes	292 (7)	101 (7)	191 (7)	
Diabetes mellitus				
No	4064 (93)	1362 (93)	2702 (92)	0.4460
Yes	323 (7)	101 (7)	222 (8)	
Chronic liver disease				
No	3652 (83)	1196 (82)	2456 (84)	0.0666
Yes	735 (17)	267 (18)	468 (16)	
Non-AIDS-defining cancers				
No	4010 (91)	1335 (91)	2675 (92)	0.8392
Yes	377 (9)	128 (9)	249 (8)	
Chronic obstructive pulmonary disease				
No	4098 (93)	1380 (94)	2718 (93)	0.0964
Yes	289 (7)	83 (6)	206 (7)	
Hypertension				
No	3826 (87)	1290 (88)	2536 (87)	0.1927
Yes	561 (13)	173 (12)	388 (13)	
Substance use disorder				
No	2667 (61)	863 (59)	1804 (62)	0.0893
Yes	1720 (39)	600 (41)	1120 (38)	
Alzheimer's/Dementia				
No	4272 (97)	1432 (98)	2840 (97)	0.1697
Yes	115 (3)	31 (2)	84 (3)	
Personality disorder				
No	3978 (91)	1317 (90)	2661 (91)	0.3160
Yes	409 (9)	146 (10)	263 (9)	
Schizophrenia				
No	4124 (94)	1377 (94)	2747 (94)	0.8707
Yes	263 (6)	86 (6)	177 (6)	
Mood/Anxiety disorder				
No	2233 (51)	767 (52)	1466 (50)	0.1620
Yes	2154 (49)	696 (48)	1458 (50)	
Rheumatoid arthritis				
No	4349 (99)	1448 (99)	2901 (99)	0.5277
Yes	38 (1)	15 (1)	23 (1)	

Continued

Table 1: Continued

Variables	Full cohort (n = 4387) n (%)	Test dataset (n = 1463) n (%)	Training dataset (n = 2924) n (%)	p-value
Metastatic solid tumor				
No	4308 (98)	1436 (98)	2872 (98)	0.9703
Yes	79 (2)	27 (2)	52 (2)	
Osteoporosis				
No	4258 (97)	1426 (98)	2832 (97)	0.2954
Yes	129 (3)	37 (2)	92 (3)	
Asthma				
No	3791 (86)	1236 (84)	2555 (87)	0.0095
Yes	596 (14)	227 (16)	369 (13)	
Peptic ulcer				
No	4079 (93)	1358 (93)	2721 (93)	0.8228
Yes	308 (7)	105 (7)	203 (7)	
Osteoarthritis				
No	4124 (94)	1370 (94)	2754 (94)	0.5179
Yes	263 (6)	93 (6)	170 (6)	
CD4 nadir, cells/mm ³				
<50	172 (4)	61 (4)	111 (4)	0.1779
50–199	590 (14)	202 (14)	388 (13)	
200–349	895 (20)	271 (19)	624 (21)	
350+	2730 (62)	929 (63)	1801 (62)	
Plasma viral load suppression				
No	1271 (29)	446 (31)	825 (28)	0.1266
Yes	3116 (71)	1017 (69)	2099 (72)	
Income assistance				
No	2799 (64)	932 (64)	1867 (64)	0.9508
Yes	1588 (36)	531 (36)	1057 (36)	
Continuous covariates	Median (Q1-Q3)	Median (Q1-Q3)	Median (Q1-Q3)	
Index year	2015 (2011-2017)	2015 (2011-2017)	2015 (2011-2017)	0.8439
Age at index date (years)	47 (38-54)	46 (38-54)	47 (39-55)	0.2169
Year of ART initiation	2009 (2006-2013)	2009 (2006-2013)	2009 (2006-2013)	0.3408
Years on ART at index date (years)	4.0 (2.0-6.0)	4.0 (2.0-6.0)	4.0 (2.0-6.0)	0.4012
CD4 nadir at index date (cells/mm ³)	430 (250-610)	426 (250-620)	430 (250-300)	0.6017
Follow-up time at end of follow-up (years)	3.3 (1.4-6.3)	3.4 (1.4-6.4)	3.2 (1.3-6.3)	0.2933

Q1-Q3: 25th - 75th percentiles; ART: antiretroviral therapy; p-value significance is set at 0.05.

the c-statistics for the training and test datasets. Model 3 demonstrated superior discrimination (0.8355, 95% CI: 0.8187–0.8523 in the training dataset and 0.7965, 95% CI: 0.7664–0.8266 in the test dataset) of one-year all-cause mortality to Models 1 and 2, and it was chosen as the final predictive model. Among all predictors in Model 3, SUD (aHR 2.60, 95% CI: 2.08–3.27), MST (aHR 2.27, 95% CI: 1.47–3.50) and NADC (aHR 2.16, 95% CI: 1.63–2.86) had the strongest association with one-year all-cause mortality. Increasing CD4 nadir per 100 cells/mm³ and years on ART were associated with a reduced aHR (aHR 0.84 per 100 cells/mm³ increase, 95% CI: 0.80–0.88 and aHR 0.94 per one-year increase, 95% CI: 0.89–0.98, respectively) for one-year all-cause mortality. Other predictors significantly associated with higher one-year all-cause mortality risk are

shown in Table 2. The unadjusted results for each variable in the training dataset are shown in Supplementary Table S3. The adjusted model (Supplementary Table S4) yielded bootstrapping-based adjusted c-statistic approximately the same as the unadjusted c-statistic and slightly lower than the c-statistic in the final predictive model, suggesting good discrimination (0.8227 [95% CI: 0.8077–0.8377] versus 0.8188 [95% CI: 0.8180–0.8196]).

MRPi interpretation and risk calculator

The median MRPi in our population was 49.6 (Q1-Q3: 43.1–57.5). As the MRPi increased, the predicted one-year all-cause mortality also increased. For example, let us consider the hypothetical case of a male participant, 46 years old, on ART

Table 2: Cox proportional models showing adjusted hazard ratios based on the training dataset and Harrell's c-statistics for the training and test datasets

Variables	Model 1 aHR (95% CI)	Model 2 aHR (95% CI)	Model 3 aHR (95% CI)
Cardiovascular disease (Ref: No)			
Yes	1.93 (1.51-2.47)	1.88 (1.45-2.42)	1.89 (1.46-2.43)
Chronic liver disease (Ref: No)			
Yes	1.79 (1.46-2.20)	1.72 (1.40-2.12)	1.72 (1.40-2.11)
Chronic kidney disease (Ref: No)			
Yes	1.85 (1.43-2.38)	1.85 (1.43-2.40)	1.62 (1.25-2.11)
Chronic obstructive pulmonary disease (Ref: No)			
Yes	2.02 (1.58-2.57)	1.84 (1.43-2.36)	2.06 (1.59-2.65)
Non-AIDS-defining cancers (Ref: No)			
Yes	2.15 (1.64-2.82)	2.00 (1.51-2.64)	2.16 (1.63-2.86)
Substance use disorder (Ref: No)			
Yes	2.92 (2.38-3.59)	3.04 (2.46-3.76)	2.60 (2.08-3.27)
Alzheimer's/Dementia (Ref: No)			
Yes	1.79 (1.30-2.47)	1.63 (1.17-2.26)	
Metastatic solid tumor (Ref: No)			
Yes	2.44 (1.59-3.76)	2.50 (1.62-3.86)	2.27 (1.47-3.50)
Diabetes mellitus (Ref: No)			
Yes			1.28 (0.95-1.75)
Hypertension (Ref: No)			
Yes		0.79 (0.61-1.03)	0.79 (0.60-1.04)
Sex at birth (Ref: Male)			
Female		1.31 (1.06-1.61)	1.22 (0.99-1.50)
Income assistance (Ref: No)			
Yes			1.33 (1.08-1.63)
^a Age per 10-year increase		1.20 (1.09-1.33)	1.35 (1.22-1.49)
Plasma viral load suppression (Ref: Yes)			
No			1.63 (1.31-2.03)
^b CD4 nadir per 100 cells/mm ³			
Linear			0.84 (0.80-0.88)
Quadratic			1.01 (1.00-1.03)
^c Years on ART			
Linear			0.94 (0.89-0.98)
Quadratic			1.02 (1.00-1.04)
Cubic			0.99 (0.99-1.00)
Dataset	Harrell's c- statistic (95% CI)	Harrell's c- statistic (95% CI)	Harrell's c- statistic (95% CI)
Training dataset	0.7766 (0.7563-0.7969)	0.7850 (0.7650-0.8050)	0.8355 (0.8187-0.8523)
Test dataset	0.7538 (0.7216-0.7859)	0.7552 (0.7228-0.7876)	0.7965 (0.7664-0.8266)

ART: Antiretroviral Therapy; aHR: adjusted hazard ratio; CI: Confidence Interval; ^aLinear form of age centered at the median (47 years); ^bPolynomial forms of CD4 nadir centered at the median (430 cells/mm³); ^cLinear form of years on ART centered at the median (4 years).

for four years with no chronic comorbidities, a CD4 nadir of 500 cells/mm³, and a suppressed pVL. This participant's estimated MRP_i was 41 and his one-year predicted all-cause mortality probability was 0.8%. If he was also diagnosed with SUD, his MRP_i would increase to 49, and his one-year predicted all-cause mortality would rise to 2.0%. Alternatively, if his CD4 nadir dropped to 150 cells/mm³, his MRP_i would be 47, and his predicted mortality would be 1.6%. Finally, if his pVL became unsuppressed, but his CD4 nadir remained the same, his MRP_i would be 45, and his predicted mortality would be 1.2%.

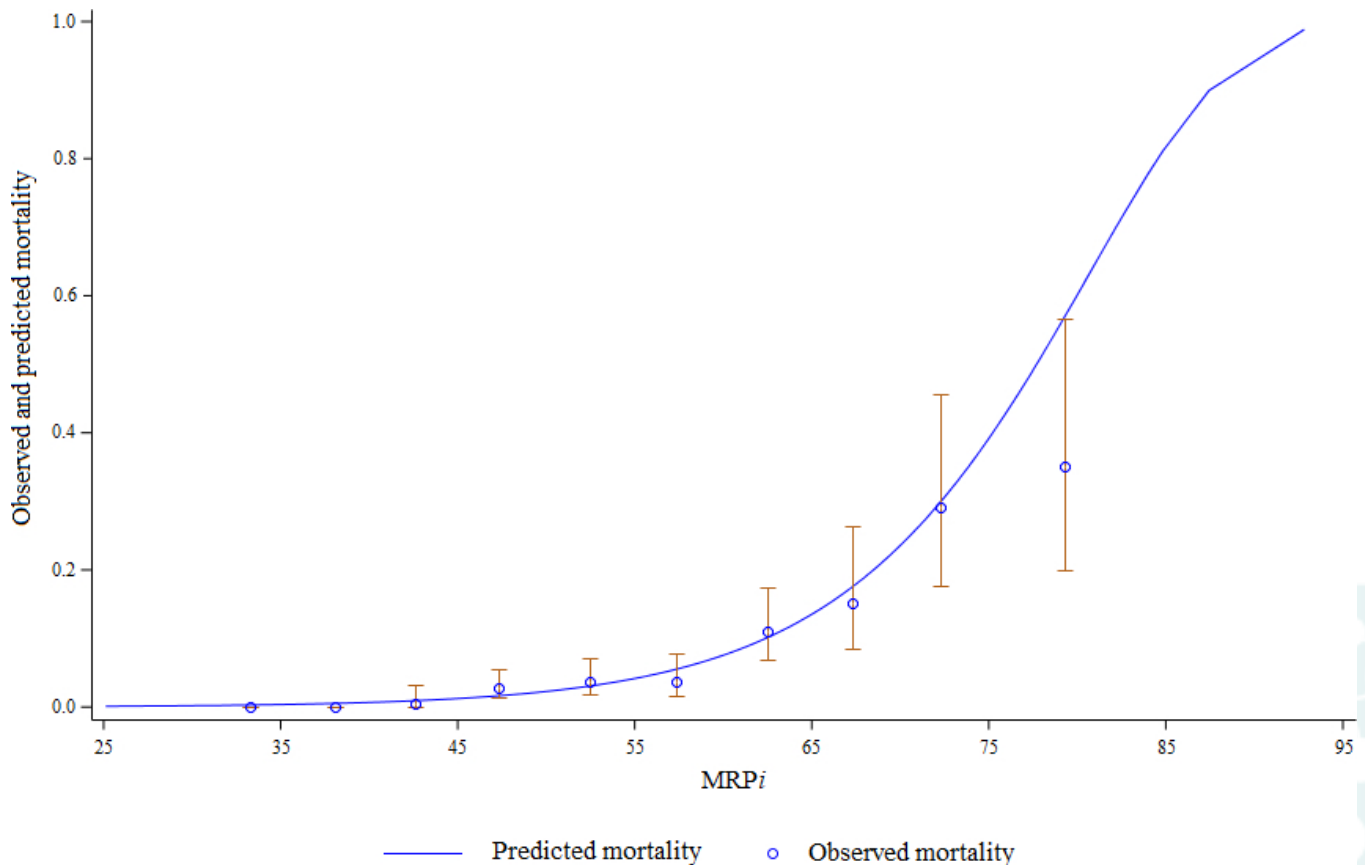
Figure 1 depicts the calibration plot comparing predicted versus observed one-year all-cause mortality probabilities across the ten subgroups of predicted mean MRP_i in the test dataset for the data-splitting internal validation method. When we applied the final Cox model to the test dataset, we found that the predicted and observed mortality were closely aligned (visually) (Supplementary Table S5) for MRP_i <75 over the one-year observation period, which includes most of the data (1,437, 98%). The model overpredicted the probability of mortality for MRP_i greater than or equal to 75. However, the less-than-perfect alignment was in the tail of the MRP_i distribution, representing those individuals who were more likely to die. Therefore, the overall calibration plot suggests the model was well calibrated with good agreement between the predicted and observed mortality probabilities. Supplementary Table S5 summarises the data in Figure 1.

Figure 2 (Supplementary Table S6) shows the calibration plot comparing the mean predicted and observed KM (unadjusted and adjusted) one-year survival probabilities using the full dataset (bootstrap-based internal validation). When the survival probability was <95%, there was a non-significant, modest difference between the predicted and observed one-year survival probabilities with overlapping 95% CI. Though there was a closer correspondence between the predicted and observed one-year survival as the one-year survival probability increased above 95%, some 95% CI did not cross the ideal line implying an underestimation of the predicted one-year survival probability in these subgroups. These inconsistencies were likely because the predicted survival probabilities in some subgroups were higher or lower depending on the sample size distribution. Note that we did not have observed survival probabilities <75% in this dataset. Overall, there was good agreement between the observed KM and predicted one-year survival probabilities suggesting a good calibration.

Secondary analyses

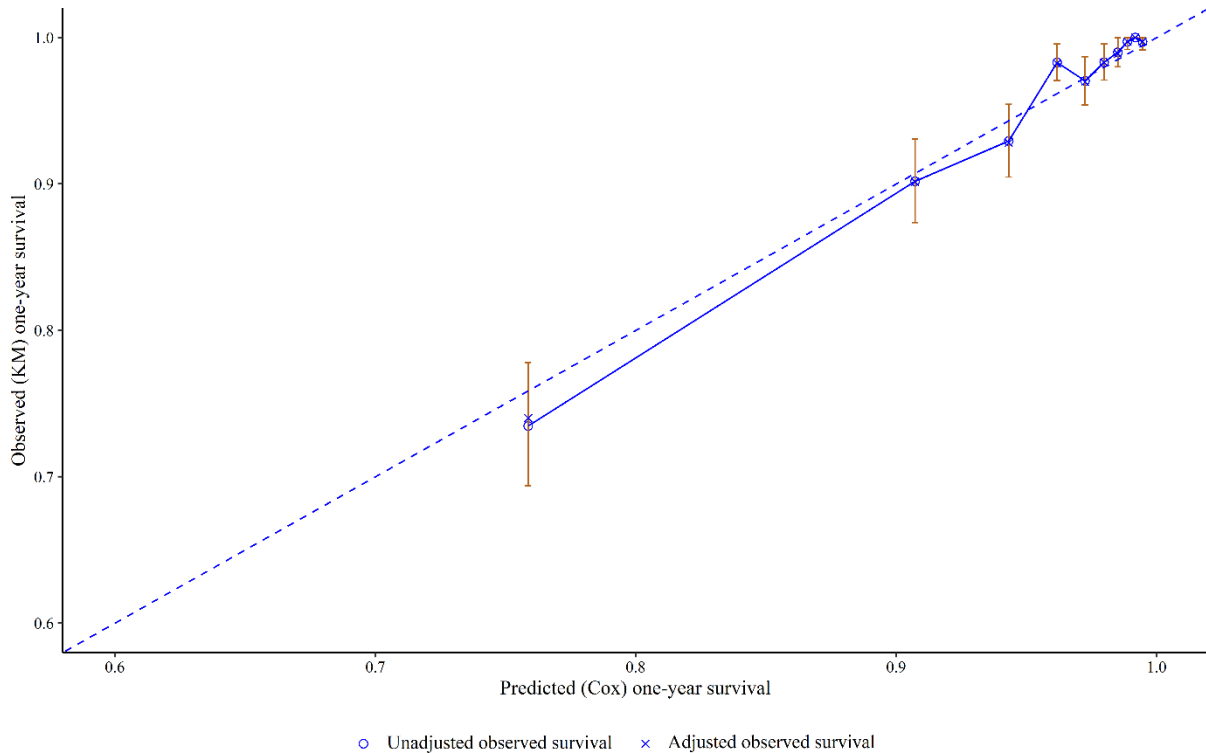
The excluded participants were significantly more likely to be younger, and have longer median years on ART and lower mortality (Supplementary Table S7). Results of the sex-based model-building are shown in Supplementary Figures S7–S10 and Tables S8–S11. Females in the final analytic sample were younger than males. About 572 (70%) females versus 2084

Figure 1: Predicted one-year all-cause mortality using our final Cox model vs observed Kaplan-Meier estimates of the one-year all-cause mortality probability by subgroups of the one-year mortality risk prediction index (MRP_i) using the test dataset



The observed one-year mortality probability is shown with 95% confidence intervals. Solid lines reflect predicted mortality probability calculated using the test dataset (n=1463). MRP_i ranges from 0-100; x-axis begins from 0-25 because the minimum score in our sample is 25.

Figure 2: Observed Kaplan-Meier estimates vs predicted estimates of one-year survival probabilities using the full dataset



KM; Kaplan-Meier. The observed one-year survival probability is shown with 95% confidence intervals. A solid line is drawn to join the two unadjusted observed survival points. The distance between a 45-degree (dotted) line with slope 1 and intercept 0 and the points reflects the difference between the observed and predicted survival probabilities. We did not have any observed survival probability <75%.

(58%) males were <50 years (Supplementary Figure S3). Sex-based analyses indicated that females tended to have a higher comorbidity burden and slightly higher MRPi scores than males. However, the overall predictive performance of the MRPi (c-statistic) remained high in both subgroups. Multimorbidity was more common among females (602, 74%), and males (958, 27%) were more likely to have no comorbidity at the index date (Supplementary Figure S4). The most prevalent comorbidities were SUD, MAD and CLD for both males and females (Supplementary Figure S5). In addition, females experienced two times the burden of SUD and CLD compared with males (66% versus 33%, and 30% versus 14%, respectively). Supplementary Figure S6 shows the distribution of the MRPi by sex at birth. A total of 473 (13%) males versus 36 (4%) females had an MRPi between 30 and 39, while 225 (28%) females versus 448 (13%) males had an MRPi between 60 and 69.

Discussion

We developed and validated an index score to predict the one-year all-cause mortality probability among PLWH using information routinely available in administrative health data. Our results showed that accounting for age, sex at birth, comorbidities, income assistance, and HIV-related markers in a single model provided the best discrimination of one-year all-cause mortality probability. By adjusting for the key variables in a single index, this study underscores the compound

effects of HIV, ART, and chronic comorbidities on all-cause mortality among the aging PLWH. In this contemporary index, comorbidities were the strongest predictors of one-year all-cause mortality among PLWH. In addition, females had a higher comorbidity burden and MRPi than males in this cohort, highlighting the need to address the drivers of their sub-optimal clinical outcomes.

The recently updated VACS 2.0 index [19], which included age, body mass index and other clinical biomarkers to estimate all-cause mortality, has been validated in large North American cohorts (c-statistic 0.819, 95% CI: 0.815-0.823) [20]. However, it contains laboratory and clinical data to identify the presence of comorbidities which are not uniformly available in clinical or administrative datasets. In such circumstances, the MRPi is an alternative index with strong discrimination (c-statistic 0.8355, 95% CI: 0.8187-0.8523). Thus, once the MRPi is externally validated, it can become a valuable new tool to characterise short-term mortality among PLWH.

The final predictors in our model represent common factors which have been shown to contribute to an increase in the health burden of PLWH [7, 10]. The presence of comorbidities such as SUD, MST, and NADC had the largest association with one-year all-cause mortality probability among all predictors assessed. SUD was also strongly associated with an increased one-year mortality probability consistent with a previous study [33]. A diagnosis of SUD was associated with drug overdose and other adverse health outcomes (e.g., high viremia and premature mortality) [34–36]. It is worth mentioning that, since 2017, BC has had record-setting

mortality rates related to the illicit drug toxicity crisis, which may further exacerbate the short-term mortality probability associated with SUD [37]. Our index was developed using data before the illicit drug toxicity crisis in BC, so our mortality risk prediction is unlikely to be biased by this crisis. Similar to previous indices, NADC increased the mortality probability in our population [13, 14, 38]. There is evidence that aging, lifestyle factors (e.g., high tobacco use) and HIV-specific factors (e.g., duration of HIV infection and immune status) increase the risk of NADC in PLWH [39, 40].

Age is a crucial social determinant of health, making our MRP_i relevant to the growing population of aging PLWH. We also found that being female was associated with increased mortality risk. Of note, females in our population were more likely to have a history of injection drug use, a higher overdose-related mortality, and they faced inequities in access and adherence to HIV care, which further compounded their mortality risk [34, 36, 41]. A high CD4 nadir at the index date was a protective factor in the MRP_i while the reverse was true for unsuppressed pVL. We have shown that by keeping all participant characteristics constant, the MRP_i and predicted mortality changed as the CD4 nadir and pVL improved or worsened indicating the direct impact of clinical management on mortality. The association between high CD4 and mortality has been attributed to improved immunologic response in PLWH [42]. Unsuppressed pVL and its association with mortality can be explained by longer exposure to inflammation, which increases the risk of chronic comorbidities and, therefore, mortality [1]. Also, unsuppressed pVL is associated with poor retention to care, which may lead to poor health outcomes [43, 44]. We found that years on ART was significantly associated with a lower mortality risk in this study, and it improved the performance of the MRP_i . ART use is associated with pVL suppression and CD4 improvement [42], which were common in over half of the study participants; it is, thus, associated with a lower mortality risk. However, our study's relatively short median time on ART may also partly explain the observed lower mortality risk. Our separate sex-based analyses confirmed that females in our cohort had a slightly higher comorbidity burden, which aligns with previous studies of aging among PLWH. While our final model performed well in both sexes, future research might explore additional sex-specific risk factors.

Our study has some potential limitations. First, administrative data are not primarily collected for use in research. Thus, to identify the comorbidities used in our study, we relied on the Canadian Chronic Disease Surveillance System case definitions, supplemented by input from epidemiology and medical care experts familiar with BC-specific claims-related policies, as well as published literature [7, 10, 24]. Second, because we used administrative data, we could not assess the clinical severity of comorbidities; however, this limitation did not prevent our index from adequately predicting mortality, which is in line with previous reports indicating that administrative data can yield models with strong discrimination and calibration [45]. Third, the median follow-up time in our study was less than five years due to selecting a random index date for each participant, which necessarily shortened the observation period for some participants. Consequently, the MRP_i is most applicable for short-term (one-year) mortality risk rather than long-term

prediction. We also note that, although participants initiated ART from 2001 to 2018, we extended follow-up through 2019, thereby capturing many modern changes in treatment practices. Nonetheless, ART continues to evolve; this could limit generalisability to individuals who started therapy more recently or outside of BC's universal healthcare setting. Fourth, although we initially sought to include ethnicity, the high proportion of missing data precluded its use; instead, we used income assistance as a proxy for socioeconomic status, which may not fully capture the full dimension of this social determinant of health. Fifth, although we considered 18 comorbidities widely recognised in the literature, there may be additional conditions relevant to PLWH that we did not capture. As more detailed or specialised data become available, future iterations of this risk index could expand the range of comorbidities included. Finally, while our use of data-splitting and bootstrapping supports the internal validity of our model, predictions for highly comorbid individuals with high mortality risk were inevitably more uncertain, given the smaller sample size in this subgroup.

Conclusions

The MRP_i provides a promising tool to assess the risk of short-term mortality among PLWH in the modern ART era. In addition, these findings highlight the need for integrated HIV care models that suit the evolving healthcare needs of aging PLWH while reflecting the intersectionality of HIV and chronic diseases.

Funding

JSGM is supported with grants paid to his institution by BC Ministry of Health, Health Canada, Canadian Institutes of Health Research, Public Health Agency of Canada, Genome Canada, Genome BC, Vancouver Coastal Health and VGH Foundation. VDL is funded by a grant from the Canadian Institutes of Health Research (PJT-148595), and the Canadian Foundation for AIDS Research (CANFAR Innovation Grant – 30-101).

Acknowledgements

The authors thank all the participants included within STOP HIV/AIDS, the British Columbia Centre for Excellence in HIV/AIDS, the BC Ministry of Health, and the institutional data stewards for granting access to the data.

Disclaimer

All inferences, opinions, and conclusions drawn in this manuscript are those of the authors and do not reflect the views or policies of the data stewards.

Conflict of interest

JSGM received institutional grants provided by Gilead Sciences Inc, Janssen, Merck Sharp & Dohme LLC, and

ViiV Healthcare. VDL received honoraria to present at the 2023 CROI (Conference on Retroviruses and Opportunistic Infections) ViiV Healthcare Ambassador Program. The other authors declare that they have no conflict of interest.

Ethics statement

This study received approval from the University of British Columbia Ethics Review Committee at the St Paul's Hospital, Providence Health Care site (H18-02208). The usage of administrative data was approved by data stewards. Due to the use of anonymised administrative data, informed consent was not required for this study.

Data availability

The British Columbia Centre for Excellence in HIV/AIDS (BC-CfE) is prohibited from making individual-level data available publicly due to provisions in our service contracts, institutional policy, and ethical requirements. To facilitate research, we make such data available via data access requests. Some BC-CfE 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-CfE researcher. For more information, please contact Mark Helberg, Senior Director, Internal and External Relations and Strategic Development: mhelberg@bccfe.ca.

Author's contribution

Concept and design: VDL; Acquisition, analysis, or interpretation of data: BTT, NF, HN, KD, JZ, JT, VDL; Drafting of the manuscript: BTT, VDL; Critical revision of the manuscript for important intellectual content: BTT, NF, HN, JZ, SE, KD, JT, KAS, RB, JSGM, VDL; Administrative, technical, or material support: RB, JSGM, VDL. All authors have read and approved the final manuscript.

References

- Deeks SG, Lewin SR, Havlir DV. The end of AIDS: HIV infection as a chronic disease. *The Lancet*. 2013;382(9903):1525-33. [https://doi.org/10.1016/S0140-6736\(13\)61809-7](https://doi.org/10.1016/S0140-6736(13)61809-7)
- Monforte AD, Sabin CA, Phillips A, Sterne J, May M, Justice A, et al. The changing incidence of AIDS events in patients receiving highly active antiretroviral therapy. *Archives of Internal Medicine*. 2005;165(4):416-23. <https://doi.org/10.1001/archinte.165.4.416>
- Teeraananchai S, Kerr S, Amin J, Ruxrungtham K, Law M. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. *HIV Medicine*. 2017;18(4):256-66. <https://doi.org/10.1111/hiv.12421>
- 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. <https://doi.org/10.1371/journal.pone.0081355>
- British Columbia Centre for Excellence in HIV/AIDS. HIV Monitoring Quarterly Report for British Columbia. Fourth Quarter 2013 2013 [Available from: <https://bccfe.ca/sites/default/files/uploads/publications/centredocs/bc-monitoring-report-13q4-updated-2015-jan-20.pdf>.
- British Columbia Centre for Excellence in HIV/AIDS. HIV Monitoring Semi-Annual Report For British Columbia. Fourth Quarter 2022 2023 [Available from: <https://stophiv aids.ca/qmr/2022-Q4/#/bc>.
- Nanditha NGA, Paiero A, Tafessu HM, St-Jean M, McLinden T, Justice AC, et al. Excess burden of age-associated comorbidities among people living with HIV in British Columbia, Canada: a population-based cohort study. *BMJ Open*. 2021;11(1):e041734. <https://doi.org/10.1136/bmjopen-2020-041734>
- Sokoya T, Steel HC, Nieuwoudt M, Rossouw TM. HIV as a cause of immune activation and immunosenescence. *Mediators of Inflammation*. 2017;2017. <https://doi.org/10.1155/2017/6825493>
- Sereti I, Altfield M. Immune activation and HIV: an enduring relationship. *Current Opinion in HIV and AIDS*. 2016;11(2):129. <https://doi.org/10.1097/COH.0000000000000244>
- Nanditha NGA, Zhu J, Wang L, Kopec J, Hogg RS, Montaner JSG, et al. Disability-adjusted life years associated with chronic comorbidities among people living with and without HIV: Estimating health burden in British Columbia, Canada. *PLOS Glob Public Health*. 2022;2(10):e0001138. <https://doi.org/10.1371/journal.pgph.0001138>
- Lachaine J, Baribeau V, Lorgeoux R, Tossonian H. Health care resource utilization and costs associated with HIV-positive patients with comorbidity versus HIV-negative patients with comorbidity. *Value in Health*. 2017;20(9):A791. <https://doi.org/10.1016/j.jval.2017.08.2323>
- Nachega JB, Hsu AJ, Uthman OA, Spinewine A, Pham PA. Antiretroviral therapy adherence and drug-drug interactions in the aging HIV population. *AIDS*. 2012;26 Suppl 1:S39-53. <https://doi.org/10.1097/QAD.0b013e32835584ea>
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases*. 1987;40(5):373-83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)

14. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Medical Care*. 1998;8-27. <https://doi.org/10.1097/00005650-199801000-00004>
15. Weiner JP, Starfield BH, Steinwachs DM, Mumford LM. Development and application of a population-oriented measure of ambulatory care case-mix. *Medical Care*. 1991;452-72. <https://doi.org/10.1097/00005650-199105000-00006>
16. Justice AC, Feinstein AR, Wells CK. A new prognostic staging system for the acquired immunodeficiency syndrome. *New England Journal of Medicine*. 1989;320(21):1388-93. <https://doi.org/10.1056/nejm198905253202106>
17. Egger M, May M, Chêne G, Phillips AN, Ledergerber B, Dabis F, et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *The Lancet*. 2002;360(9327):119-29. [https://doi.org/10.1016/s0140-6736\(02\)09411-4](https://doi.org/10.1016/s0140-6736(02)09411-4)
18. Justice AC, McGinnis K, Skanderson M, Chang C, Gibert C, Goetz M, et al. Towards a combined prognostic index for survival in HIV infection: the role of 'non-HIV' biomarkers. *HIV Medicine*. 2010;11(2):143-51. <https://doi.org/10.1111/j.1468-1293.2009.00757.x>
19. Tate JP, Sterne JA, Justice AC, Study VAC, Collaboration ATC. Albumin, white blood cell count, and body mass index improve discrimination of mortality in HIV-positive individuals. *AIDS (London, England)*. 2019;33(5):903. <https://doi.org/10.1097/qad.0000000000002140>
20. McGinnis KA, Justice AC, Moore RD, Silverberg MJ, Althoff KN, Karris M, et al. Discrimination and Calibration of the Veterans Aging Cohort Study Index 2.0 for Predicting Mortality Among People With Human Immunodeficiency Virus in North America. *Clinical Infectious Diseases*. 2022;75(2):297-304. <https://doi.org/10.1093/cid/ciab883>
21. Justice AC, Modur S, Tate JP, Althoff KN, Jacobson LP, Gebo K, et al. Predictive accuracy of the Veterans Aging Cohort Study (VACS) index for mortality with HIV infection: a North American cross cohort analysis. *Journal of Acquired Immune Deficiency Syndromes (1999)*. 2013;62(2):149. <https://doi.org/10.1097/QAI.0b013e31827df36c>
22. Qiao S, Li X, Olatosi B, Young SD. Utilizing Big Data analytics and electronic health record data in HIV prevention, treatment, and care research: a literature review. *AIDS Care*. 2021;1-21. <https://doi.org/10.1080/09540121.2021.1948499>
23. 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>
24. Nanditha NGA, Dong X, McLinden T, Sereda P, Kopec J, Hogg RS, et al. The impact of lookback windows on the prevalence and incidence of chronic diseases among people living with HIV: an exploration in administrative health data in Canada. *BMC Medical Research Methodology*. 2022;22(1):1. <https://doi.org/10.1186/s12874-021-01448-x>
25. BC Ministry of Health. [creator] 2017. Vital Statistics Deaths. BC Ministry of Health [publisher]. Data extract. BC Vital Statistics Agency (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
26. McGinnis KA, Justice AC, Marconi VC, Rodriguez-Barradas MC, Hauser RG, Oursler KK, et al. Combining Charlson comorbidity and VACS indices improves prognostic accuracy for all-cause mortality for patients with and without HIV in the Veterans Health Administration. *Front Med (Lausanne)*. 2023;10:1342466. <https://doi.org/10.3389/fmed.2024.1532350>
27. British Columbia PharmaCare for health professionals. Pharmacies. Product Identification Numbers (PINS) 2023 [Available from: <https://www2.gov.bc.ca/gov/content/health/practitioner-professional-resources/pharmacare/pharmacies/product-identification-numbers>].
28. McDonald JH. Handbook of biological statistics. Baltimore (US): Sparky House Publishing; 2009.
29. Harrell Jr FE. Regression Modeling Strategies [Internet]. 2nd edition. New York (US): Springer-Verlag; 2015. [cited 2023 March 20].
30. 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>
31. Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*. 1996;15(4):361-87. [https://doi.org/10.1002/\(SICI\)1097-0258\(19960229\)15:4<361::AID-SIM168>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4)
32. Pencina MJ, D'Agostino RB, Sr. Evaluating Discrimination of Risk Prediction Models: The C Statistic. *JAMA*. 2015;314(10):1063-4. <https://doi.org/10.1001/jama.2015.11082>
33. Nigussie F, Alamer A, Mengistu Z, Tachbele E. Survival and Predictors of Mortality Among Adult HIV/AIDS Patients Initiating Highly Active Antiretroviral Therapy in Debre-Berhan Referral Hospital, Amhara, Ethiopia: A Retrospective Study. *HIV AIDS (Auckl)*. 2020;12:757-68. <https://doi.org/10.2147/HIV.S274747>

34. Salters KA, Parent S, Nicholson V, Wang L, Sereda P, Pakhomova TE, et al. The opioid crisis is driving mortality among under-served people living with HIV in British Columbia, Canada. *BMC public health*. 2021;21(1):1-8. <https://doi.org/10.1186/s12889-021-10714-y>
35. Cohn SE, Jiang H, McCutchan JA, Koletar SL, Murphy RL, Robertson KR, et al. Association of ongoing drug and alcohol use with non-adherence to antiretroviral therapy and higher risk of AIDS and death: results from ACTG 362. *AIDS Care*. 2011; 23(6):775-85. <https://doi.org/10.1080/09540121.2010.525617>
36. St-Jean M, Dong X, Tafessu H, Moore D, Honer WG, Vila-Rodriguez F, et al. Overdose mortality is reducing the gains in life expectancy of antiretroviral-treated people living with HIV in British Columbia, Canada. *International Journal of Drug Policy*. 2021;96:103195. <https://doi.org/10.1016/j.drugpo.2021.103195>
37. British Columbia Coroners Service. Illicit drug toxicity deaths in BC: January 1, 2012–August 31, 2022. Vancouver: Ministry of Public Safety & Solicitor General; 2022.
38. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *American Journal of Epidemiology*. 2011;173(6):676-82. <https://doi.org/10.1093/aje/kwq433>
39. Deeken JF, Tjen ALA, Rudek MA, Okuliar C, Young M, Little RF, et al. The rising challenge of non-AIDS-defining cancers in HIV-infected patients. *Clin Infect Dis*. 2012;55(9):1228-35. <https://doi.org/10.1093/cid/cis613>
40. Chiao EY, Coghill A, Kizub D, Fink V, Ndlovu N, Mazul A, et al. The effect of non-AIDS-defining cancers on people living with HIV. *The Lancet Oncology*. 2021;22(6):e240-e53. [https://doi.org/10.1016/S1470-2045\(21\)00137-6](https://doi.org/10.1016/S1470-2045(21)00137-6)
41. Carter A, Min JE, Chau W, Lima VD, Kestler M, Pick N, et al. Gender inequities in quality of care among HIV-positive individuals initiating antiretroviral treatment in British Columbia, Canada (2000-2010). *PLoS One*. 2014;9(3):e92334. <https://doi.org/10.1371/journal.pone.0092334>
42. Moore DM, Harris R, Lima V, Hogg B, May M, Yip B, et al. Effect of baseline CD4 cell counts on the clinical significance of short-term immunologic response to antiretroviral therapy in individuals with virologic suppression. *J Acquir Immune Defic Syndr*. 2009;52(3):357-63. <https://doi.org/10.1097/QAI.0b013e3181b62933>
43. Teixeira da Silva DS, Luz PM, Lake JE, Cardoso SW, Ribeiro S, Moreira RI, et al. Poor retention in early care increases risk of mortality in a Brazilian HIV-infected clinical cohort. *AIDS Care*. 2017;29(2):263-7. <https://doi.org/10.1080/09540121.2016.1211610>
44. Yehia BR, French B, Fleishman JA, Metlay JP, Berry SA, Korthuis PT, et al. Retention in care is more strongly associated with viral suppression in HIV-infected patients with lower versus higher CD4 counts. *J Acquir Immune Defic Syndr*. 2014;65(3):333-9. <https://doi.org/10.1097/QAI.0000000000000023>
45. Rothberg MB, Pekow PS, Priya A, Zilberberg MD, Belforti R, Skiest D, et al. Using highly detailed administrative data to predict pneumonia mortality. *PLoS ONE*. 2014;9(1):e87382. <https://doi.org/10.1371/journal.pone.0087382>



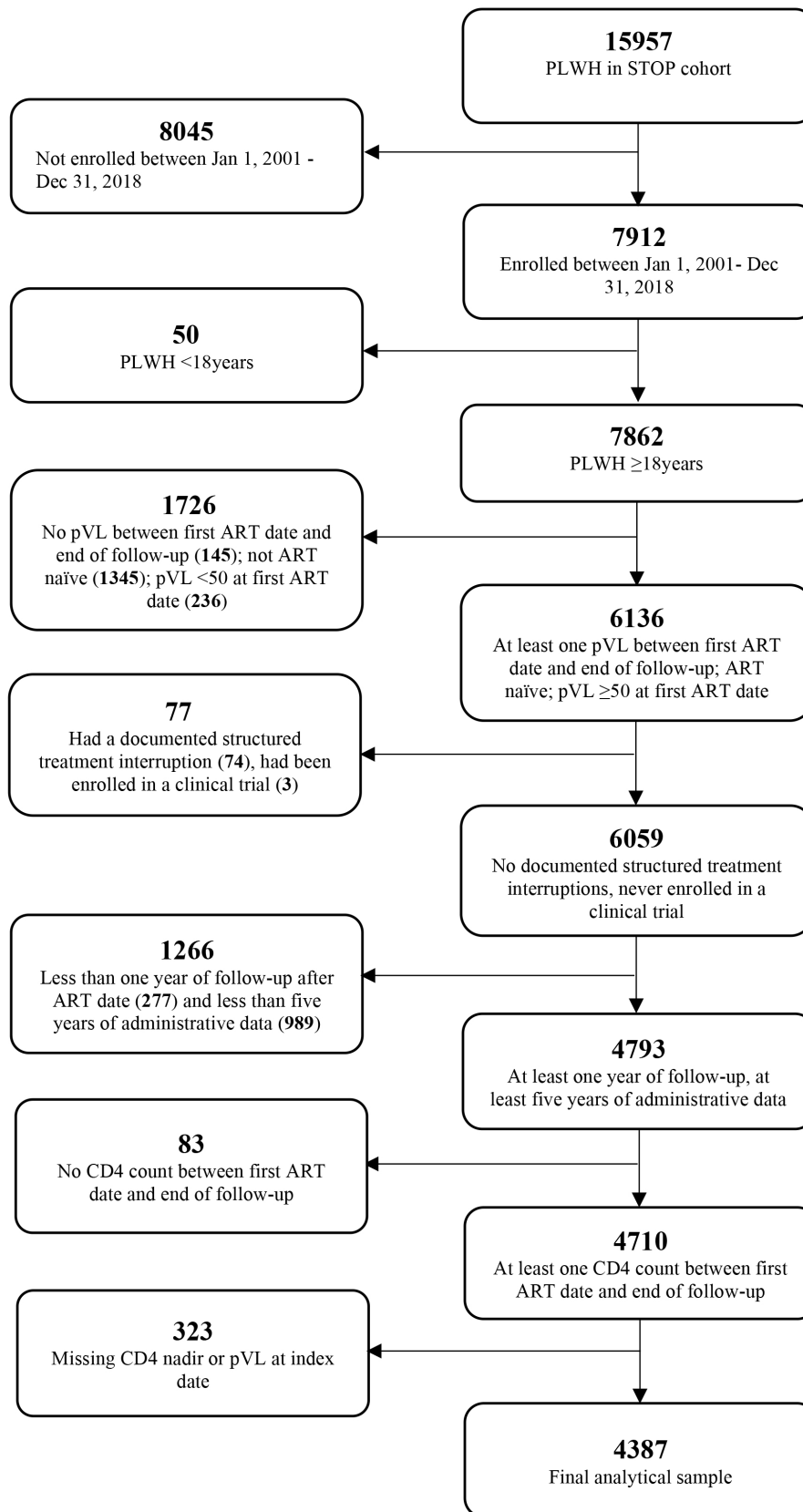
Supplementary material

Supplementary Table S1: List of linked datasets and stewards from the Seek and Treat for Optimal Prevention of HIV/AIDS (STOP HIV/AIDS) population-based cohort utilised in this study

Data Steward	Database	Description and Key data
British Columbia Centre for Excellence in HIV/AIDS (BCCfE)	Drug Treatment and Laboratory Registry	This database captures clinical and laboratory tests and HIV-related treatment and treatment outcomes. It includes data on antiretroviral (ARV) therapy history, approximately 85% of CD4 cell measurements, plasma viral load testing, the occurrence of AIDS-defining illnesses, drug resistance testing and key patient demographic data [1, 2]. In British Columbia, ART and related medical and laboratory monitoring are available free of charge to all people living with HIV residing in the province.
British Columbia Centre for Disease Control (BCCDC)	Provincial HIV/AIDS Database	This database collates all HIV testing events in BC. It includes positive HIV tests, the date of the first positive test, and the occurrence of AIDS-defining illness data [3].
British Columbia Ministry of Health	Medical Services Plan (MSP) billing database	Captures HIV and non-HIV-related outpatient and inpatient care billing for all medical services provided as fee-for-service. Also contains data on diagnostic and laboratory procedures, encounter claims and all fee-for-service dates, diagnoses and costs billed [4].
	PharmaNet	A real-time system that captures all prescriptions for drugs and medical supplies dispensed from community pharmacies and hospital outpatient pharmacies. Data include drug type and name, date and length of prescription, dosage, prescriber code and drug costs. This database does not capture ARV dispensing data [5].
	The Client Roster or Consolidation File	This database captures individual demographic and geographic data. This database is also used to construct population denominators [6].
	Discharge Abstract Database (DAD)	This database captures all discharges, transfers and deaths of all in-patients and day surgery patients from all acute care hospitals across BC [7].
	Vital Statistics	This dataset records death including the date and probable causes of death for all BC Residents [8].

Note: Heath et al., 2014 [9] and Nosyk et al., 2013 [10] provided additional details describing these data sources.

Supplementary Figure S1: Flowchart outlining the derivation of final analytical sample of 4387 people living with HIV in British Columbia, Canada, from 2002–2019



Note: PLWH: People living with HIV; STOP: Seek and Treat for Optimal Prevention of HIV/AIDS; pVL: Viral load; ART: Antiretroviral therapy.

Supplementary Table S2: List of selected comorbidities and relevant case-finding algorithms

Outcome	Case definitions	Diagnostic, Procedure, Drug code	References
Chronic liver disease (CLD)	Cirrhosis	ICD-9 for physician visit: 571	[11]
	1 hospitalisation with a cirrhosis diagnostic code	ICD-9 for hospitalisation: 456.1, 571.2, 571.5	
	OR	ICD-10: I85.9, I98.2, K70.3, K71.7, K74.6	
	2 physician visits with cirrhosis diagnostic code(s) in one year		[11]
	Decompensated Liver Cirrhosis	ICD-9: 456.0, 456.2, 572.2, 572.3, 572.4, 782.4, 789.5:	
	1 hospitalisation with a with a decompensated cirrhosis diagnostic code	ICD-10: I85.0, I86.4, I98.20, I98.3, K72.1, K72.9, K76.6, K76.7, R17, R18	
Chronic liver disease (CLD)	OR		[12]
	2 physician visits with decompensated cirrhosis diagnostic code(s) in one year		
	Chronic liver disease (CLD)	ICD-9: 070.2, 070.3, 070.41, 070.44, 070.51, 070.54, 070.7, 275.0, 573.3	
	1 hospitalisation with a CLD diagnostic code	ICD-10: B16, B17.10, B17.11, B18.0, B18.1, B18.2, B19.10, B19.11, B19.20, B19.21, E83.11, K71.6, K75.9	[13]
	OR		
	2 physician visits with CLD diagnostic code(s) in one year		
Alzheimer's/Dementia (Applicable to persons aged 40 years and older)	1 hospitalisation with an Alzheimer's/dementia diagnostic code	ICD-9: 290, 294, 331, 331.0, 331.1, 331.5	[13]
	OR	ICD-10: G30, G31.02, F00, F01, F02, F03	
	OR	DIN List	
	1 prescription for Alzheimer/dementia	2419866, 2419874, 2397595, 2397609, 2397617, 2397625, 2324059, 2324067, 2432684, 2432692, 2362260, 2362279, 2293021, 2293048, 2293056, 2366487, 2336715, 2336723, 2336731, 2336758, 2232043, 2232044, 2269457, 2269465, 2400561, 2400588, 2425157, 2425165, 2425173, 2427567, 2427575, 2427583, 2427591, 2376334, 2412853, 2412861, 2295245, 2295229, 2295237, 2447002, 2447010, 2307685, 2307693, 2307707, 2307715, 2416417, 2416425, 2420597, 2420600, 2421453, 2421461, 2395584, 2395592, 2402645, 2402653, 2413671, 2413698, 2260638, 2425343, 2425351, 2242116, 2242117, 2242118, 2245240, 2242115, 2302853, 2432803, 2302845, 2416573, 2416581, 2416603, 2443015, 2443023, 2443031, 2386011, 2386003, 2386046, 2386038, 2398885, 2398893, 2419238, 2419246, 2419254, 2404419, 2404427, 2416948, 2416956, 2402092, 2402106, 2420821, 2420848, 2420856, 2409887, 2409895, 2401614, 2401622, 2401630, 2401649, 2349116, 2443082, 2446049, 2425742, 2408600, 2408619, 2406985, 2406993, 2407000, 2407019, 2359472, 2359480, 2339439, 2339447, 2339455, 2430371, 2332809, 2332817, 2332825, 2332833, 2439557, 2439565, 2321602, 2321599, 2305984, 2305992, 2306018, 2306026, 2316943, 2316951, 2316978, 2427273, 2308169, 2308177, 2308185, 2308193, 2322331, 2322358, 2333376, 2333384, 2333392, 2398370, 2398389, 2398397, 2321130, 2382830, 2306034, 2306042, 2306050, 2306069, 2446669, 2446677, 2381508, 2381516, 2421364, 2320908, 2311283, 2311291, 2311305, 2311313, 2244300, 2244298, 2244299, 2244302, 2266717, 2266725, 2266733, 2412918, 2412934, 2348950, 2416999, 2417006, 2417014, 2417022, 2312492, 2312506, 2312514, 2423413, 2312522, 2423421, 2375729, 2375737, 2375745, 2375753, 2328666, 2328682, 2367688, 2367696, 2392283, 2392291, 2392305, 2344807, 2375532, 2324563, 2324571, 2324598, 2324601, 2426307, 2426293, 2428482, 2428490, 2340607, 2340615, 2377950, 2377969, 2377977, 2426943, 2426951	[14, 15]
	OR		
	3 physician visits with Alzheimer/dementia within 2 years but interspersed by at least 30 days.		
Diabetes mellitus (DM)	1 hospitalisation with a DM diagnostic code	ICD-9: 250	[14, 15]
	OR	ICD-10: E10, E11, E13, E14	

Continued

Supplementary Table S2: Continued

Outcome	Case definitions	Diagnostic, Procedure, Drug code	References
	2 physician visits in one year with DM diagnostic code(s)	Exclusion: Cases of suspected gestational diabetes in women aged 10-54 are not included by excluding hospitalizations, physician claims or prescriptions within the period 120 days preceding or 180 days after hospital records containing birth-related diagnostic codes.	
	OR		
	2 or more insulin prescriptions in one year	ICD-9 excluded: (641, 642, 643, 644, 645, 646, 647, 648, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 66, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 763, V27)	
	OR		
	2 or more oral antihyperglycemic (not including metformin) prescriptions in one year	ICD-10 excluded: (V27, O10, O11, O12, O13, O14, O15, O16, O21, O22, O23, O24, O25, O26, O28, O29, O30, O31, O32, O33, O34, O35, O36, O37, O40, O41, O42, O43, O44, O45, O46, O47, O48, O60, O61, O62, O63, O64, O65, O66, O67, O68, O69, O70, O71, O72, O73, O74, O75, O85, O86, O87, O88, O89, O90, O91, O92, O94, O95, O98, O99, Z37).	
	OR		
	1 insulin and one oral antihyperglycemic (including metformin) in one year	DIN:	
	OR	Antihyperglycemic:	
	2 metformin prescriptions and one physician visit with DM diagnostic code(s) in one year.	2303442, 2303450, 2303469, 2339587, 2339595, 2391600, 2429764, 2429772, 2302861, 2302888, 2302896, 2321475, 2321483, 2321491, 2423286, 2242572, 2242573, 2242574, 2374587, 2374595, 2245272, 2245273, 2245274, 399302, 312711, 1913654, 1913662, 2245247, 2297795, 2407124, 2405067, 2295377, 2295385, 2295393, 2302942, 2302950, 2302977, 2355663, 2355671, 2355698, 2403366, 2326477, 2326485, 2326493, 13889, 2229519, 2438658, 2449765, 808733, 808741, 2298279, 2298287, 2298295, 2373270, 2373289, 2373297, 2354349, 2354357, 2354365, 2417189, 2417197, 2417200, 21849, 2273756, 2403374, 2403382, 2348578, 2340763, 2340771, 2384906, 2384914, 2384922, 2424258, 2424266, 2424274, 2363518, 2363704, 2363712, 2247085, 2247086, 2247087, 2248440, 2248441, 2258781, 2258803, 2258811, 2241112, 2241113, 2241114, 2241111, 2363232, 2363240, 2273764, 2273772, 21350, 2363259, 2444844, 2444852, 2448610, 2361809, 2361817, 271330, 12556, 12564, 586773, 156728, 209937, 377937, 430986, 2274248, 2274256, 2274264, 2274272, 2224550, 2224569, 12599, 454753, 1987836, 1987534, 24708, 24716, 765996, 2356422, 2242987, 15598, 2234513, 2234514, 2345854, 2345862, 2336316, 2020734, 2020742, 2333554, 2375842, 237000, 12602, 12610, 1987828, 2224798, 1987542, 2224771, 2419300, 2419319, 2419327, 2419335, 2419343, 2419351, 2229595, 2229596, 2307669, 2307677, 2307723, 2307170, 2307634, 2307642, 2307650, 2307553, 2307561, 2307588, 2421674, 2421682, 2421690, 2443635, 2443643, 720933, 720941, 480304, 244449, 420336, 480290, 2435462, 2435470, 2279061, 2279088, 2279126, 2248210, 2287072, 2155850, 2248453, 2190885, 2190893, 2239924, 2239925, 2307189, 2307197, 2345366, 2345374, 2345382, 2417049, 2417057, 2417065, 2374013, 2374021, 2374048, 2294400, 2284545, 2284553, 2313596, 2236733, 2236734, 2303124, 2303132, 2303140, 22303140, 2354926, 2354934, 2354942, 2306166, 2306174, 2306182, 2357887, 2357895, 2357909, 2357917, 2357925, 2239926, 2230444, 2230443, 2350459, 2350467, 2147521, 2226804, 2147548, 2226812, 2230036, 2230037, 1959352, 1959360, 2238471, 2239476, 2238469, 2239474, 2238470, 2239475, 2236985, 2238698, 2236986, 2316544, 2312050, 2312069, 2312077, 2439328, 2375850, 2375869, 2375877, 2389312, 2389290, 2389304, 2273136, 2273101, 2273128, 1900927, 1900935, 2301423, 2301431, 2301458, 2415968, 2415976, 2415984, 2231095, 2231096, 2237531, 2236543, 2236548, 2242095, 2242096, 2254719, 2269589, 2269597, 2269600, 2269619, 2248008, 2248009, 2297906, 2297914, 2297922, 2357453, 2357461, 2357488, 2326329, 2326337, 2326345, 2455404, 2455412, 2455420, 2455439, 2455447, 2455455, 2425483, 2425491, 2365529, 2365537, 2397307, 2437899, 13730, 2245439, 2245440, 2245438, 2456575, 2456583, 2456591, 2456605, 2456613, 2456621, 2238103, 1913670, 1913689, 2274914, 2274922, 2274930, 2366347, 2366355, 2366363, 2354144, 2333856, 2333864, 2333872, 2416786, 2416794, 2416808, 2303922, 2388839, 2388847, 2443937, 2443945, 2403250, 2403269, 2403277, 2417219, 2417227, 2417235, 2389169, 2389177, 2389185, 2084341, 2354152, 2354160, 312762, 156663, 431168, 93033, 178543, 209872, 2418002, 2418010, 2418029, 2370921, 2228920, 2228939, 2448599, 2448602, 2434121, 2434148, 2434156, 2351064, 2351056, 2085887, 2449935, 2449943, 2320754, 2320762, 2320770	

Supplementary Table S2: Continued

Outcome	Case definitions	Diagnostic, Procedure, Drug code	References
		Insulin: 612162, 2279460, 2279479, 2279487, 2294346, 1985973, 773654, 1986821, 2022230, 632694, 632678, 1985965, 612359, 650935, 2024292, 2024306, 2024217, 2025248, 2024314, 2024322, 2024241, 2024225, 2024268, 2024233, 2024284, 2024276, 1986791, 612197, 644358, 612200, 2024446, 2024403, 2025256, 612189, 2265435, 2265443, 2244353, 2245397, 2377209, 274127, 542938, 1934066, 612235, 542946, 514551, 2229704, 2229705, 2403412, 2233562, 274119, 539201, 612219, 612170, 2439611, 2240294, 2403420, 2240295, 2240297, 2403439, 2241283, 1962639, 889113, 1962655, 889105, 795879, 1959212, 1962647, 889091, 1962663, 889121, 646148, 587737, 1959239, 2403447, 2241310, 586714, 2415089, 1959220, 733075, 2275864, 2275872, 66123203, 446564, 446580, 446572, 446610, 446602, 513644, 446599, 614416, 1985981, 1985949, 2022249, 1986813, 632651, 552275, 1985957, 552259, 999717, 5894, 1934112, 546348, 612227, 539244, 552267, 1934082, 612251, 275417, 554820, 6009, 648094, 1934074, 2245689, 2251930, 2276410, 2294338, 542911, 1934090, 612278, 275409, 514535, 2412829, 2271842, 2441829, 1934104, 612243, 275425, 1985930, 1986085, 632686, 1986805, 628301, 723789, Metformin: 2365286, 2408228, 2365294, 2408236, 2409283, 2409291, 2257726, 2257734, 2246965, 2284782, 2284790, 2343606, 2343614, 2353377, 2353385, 2378841, 2378868, 2334437, 2238827, 2334445, 2230027, 2385341, 2385368, 2406020, 2406039, 2230026, 2220628, 2231058, 2388766, 2167786, 2229785, 2305062, 2388774, 2148765, 2229656, 2341603, 2341522, 2438275, 2438283, 2364506, 2364514, 2265575, 2265583, 2361264, 2361272, 2162822, 2229517, 2378116, 2378124, 2246964, 2229994, 2242726, 2421828, 2421836, 2242589, 2223562, 2099233, 2162849, 2446065, 314552, 2268493, 2300451, 2268507, 2229516, 2239214, 2252945, 2252953, 2449390, 2449404, 2314894, 2314908, 2339110, 2339129, 2269031, 2269058, 2242931, 2242974, 2233999, 2239081, 2242783, 2246820, 2246821, 2378043, 2378051, 2380196, 2380218, 2380722, 2380730, 2379767, 2379775, 2331519, 2331527, 2045710, 2230475, 2378620, 2378639, 2188902, 2444933, 2444941, 2230671, 2230670, 2231389, 2242793, 2242794.	
Hypertension (Applicable to persons aged 20 years and older)	1 hospitalisation with a hypertension diagnostic code OR 2 physician visits with hypertension diagnostic code(s) in one year.	ICD-9: 401, 402, 403, 404, 405 ICD-10: I10, I11, I12, I13, I15	[16]
Cardiovascular Disease (CVD)	Ischemic Heart Disease (IHD) (Applicable to persons aged 20 years and older) 2 physician visits with Angina ICD-9 code 413 plus 1 prescription in one year; OR 1 specialist visit with Angina ICD-9 code 413 plus 1 prescription in one year; OR 2 physician visits with two ICD-9 codes 410, 411, 412, 413, and 414 in one year; OR 1 CABG, PCI/PCTA procedure code OR	ICD-9: 410, 411, 412, 413, 414 ICD-10: I20, I21, I22, I23, I24, I25 Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures (CCP): *Coronary artery bypass surgery (CABG): 4811, 4812, 4813, 4814, 4815, 4816, 4817, 4819. *Percutaneous Coronary Intervention/Percutaneous Transluminal Coronary Angioplasty (PCI/PTCA): 4802, 4803 Canadian Classification of Health Interventions (CCI): *CABG: 1IJ57LA, 1IJ57VS, 1IJ76. *PCI/PTCA: 1IJ50, 1IJ57G. DIN: 3662 14214 15954 15962 37613 37621 104736 125849 202959 202967 202975 208973 243116 279536 299790 342734 441686 441694 442925 446661 446688 446696 458686 458694 476579 476587 476595 476609 525529 576174 584223 584258 584266 590673 590681 590738 590746 602892 608785 614262 658812 660655 670944 677469 695726 725242 740721 749362 749370 749389 749397 754188 778869 786640 786667 786683 852384	[17]

Continued

Supplementary Table S2: Continued

Outcome	Case definitions	Diagnostic, Procedure, Drug code	References
Chronic kidney disease	1 hospitalisation with any IHD code(s).	860778 874213 874248 1911902 1911910 1911929 1913921 1926454 1926721 1927809 1946196 2011271 2041715 2041723 2042606 2042614 2042622 2044811 2046156 204616 2058472 2125218 2126559 2145294 2145308 2145316 2162806 2163527 2163535 2164337 2182734 2213370 2230732 2230733 2230734 2231441 2238998 2243588 2272830 2301288 2393433 2407442 2407450 2407469 2407477	
	Transient Ischemic Attack (TIA)	ICD-9: 435	[18]
	(Applicable to persons aged 20 years and older)	ICD-10: H34.0, G45.0, G45.1, G45.2, G45.3, G45.8, G45.9	
	1 hospitalisation with a TIA diagnostic code.		
	Cerebrovascular Accident (CVA/Stroke):	ICD-9: 362.3, 430, 431, 433.x1, 434.x, 436	[19]
	(Applicable to persons aged 20 years and older)	ICD-10: H34.1, I60, I61, I63, I64	
	1 hospitalisation with a CVA diagnostic code.	Excluding I63.6 (Cerebral infarction due to cerebral venous thrombosis) Exclusions: any traumatic brain injury (ICD-9: 800, 801, 802, 803, 804, 850, 851, 852, 853, 854, V57.x ICD-10: S02.x, S02.5, S06.x, Z50.x)	
	Congestive Heart Failure (CHF)	ICD-9: 428	[20]
	1 hospitalisation with a CHF diagnostic code	ICD-10: I50	
	OR		
	2 physician visits in one year with CHF diagnostic code(s)		
	Acute Myocardial Infarction (AMI)	ICD-9: 410	[21]
	1 hospitalisation with an AMI diagnostic code.	ICD-10: I21, I22	
	Chronic kidney disease (CKD)	ICD-9: 581, 582, 583, 585, 586, 587, 589	[22]
	1 hospitalisation with a CKD diagnostic code	ICD-10: N01, N03, N04, N05, N06, N07, N18, N19, N26, N27	
Chronic obstructive pulmonary disease (COPD) (Applicable to person 35 years and older)	OR		
	2 physician visits in one year with CKD diagnostic code(s)		
	Dialysis	00308, 00323, 00350, 00351, 00352, 00355, 00356, 00358, 00359, 00361, 00390, 33708, 33723, 33750, 33751, 33752, 33755, 33756, 33758, 33759, 33761, 33790	[23]
	9 or more physician visits within 90 days with fee item code(s)		
	Kidney Transplant	CCI: 675	[24]
Non-AIDS-defining cancers (NADC)	1 hospitalisation for kidney transplant with procedure code(s)	CCP: 1PC85	
	1 hospitalisation with COPD diagnostic code	ICD-9: 491, 492, 496	[25]
	OR	ICD-10: J41, J42, J43, J44	
Non-AIDS-defining cancers (NADC)	2 physician visits in one year with COPD diagnostic code(s)		
	1 hospitalisation with cancer diagnostic code	ICD-9: 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 170, 171, 172, 173, 174, 175, 179, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239	[26]
	OR		
	2 physician visits in 1 years with cancer diagnostic code(s)		

Continued

Supplementary Table S2: Continued

Outcome	Case definitions	Diagnostic, Procedure, Drug code	References
		ICD-10: C00,C01- C02, C03, C04, C05, C06, C07 - C08, C09, C12, C13,C14, C10, C11,C15, C16, C17, C18, C26.0, C19, C20, C21, C22.0, C22.1, C22.2, C22.3, C22.4, C22.7, C22.9, C23, C24, C25, C26.0, C26.2 - C26.9, C30, C31, C32, C33, C34, C38.0, C38.1, C38.2, C38.3, C38.4, C38.8, C39, C40 - C41, C43, C45, C48, C47, C49, C50, C51, C52, C57, C58, C60, C63, C54, C55, C56, C61, C62, C64, C65, C66, C68, C67, C69, C70, C71, C72, C73, C37, C74 - C75, C77-C80, C81, C90.0, C90.2, C91 - C95, C90.1, C26.1, C44, C76-C80, C88, C96.0-C96.2, C96.7-C96.9, C97	
Substance use disorder (SUD)	1 hospitalisation with a SUD diagnostic code OR 2 physician visits in one year with SUD diagnostic codes, followed by further condition-related codes in subsequent years.	ICD-9: 291, 292, 303, 304, 305 ICD-10: F10, F11, F12, F13, F14, F15, F16, F17, F18, F19	[27]

Note: ICD 9: International Classification of Diseases (Ninth Revisions); ICD 10: International Classification of Diseases (Tenth Revisions). CLD and CKD may include acute kidney or liver conditions unspecified by the ICD codes.



Supplementary Text S1. Creation of mortality risk prediction index scores and one-year all-cause mortality prediction.

Step 1:

We applied our final predictive Cox proportional hazard model to the training dataset and estimated the regression coefficients in the model.

The predicted one-year all-cause mortality can be derived using the conditional survival function $S(t|x) = So(t)^{\exp(x^T \beta)}$, where $So(t)$ is the baseline survival function, x^T is a vector of covariates and β is a vector of coefficients corresponding to the covariates.

Thus, the predicted one-year all-cause mortality can be shown as:

$$\begin{aligned} 1 - \text{Predicted one-year all-cause mortality} &= So(t)^{\exp(x^T \beta)} \\ \text{Predicted one-year all-cause mortality} &= 1 - So(t)^{\exp(x^T \beta)} \\ &= 1 - 0.9912^{\exp(x^T \beta)} \end{aligned}$$

Where,

$$\begin{aligned} x^T \beta &= \sum_{i=1}^{18} X_{i\beta_i} \\ &= CVD \times 0.6350 + CLD \times 0.5407 + CKD \times 0.4836 + COPD \times 0.7202 + NADC \times 0.7688 + SUD \times 0.9571 \\ &\quad + HTN \times (-0.2402) + MST \times 0.8199 + DM \times 0.2500 + IA \times 0.2818 + AGE \text{ LINEAR} \times 0.2966 + SEX \times 0.1976 \\ &\quad + SUPPRESSION \times 0.4879 + CD4 \text{ LINEAR} \times (-0.1807) + CD4 \text{ QUADRATIC} \times 0.0135 + YEARS \text{ ON ART LINEAR} \\ &\quad \times (-0.0664) + YEARSON \text{ ART QUADRATIC} \times 0.0177 + YEARS \text{ ON ART CUBIC} \times (-0.0019) \end{aligned}$$

Note: ART: Antiretroviral therapy; CVD: Cardiovascular disease; CLD: Chronic liver disease; CKD: Chronic kidney disease; COPD: Chronic obstructive pulmonary disease; NADC: Non-AIDS-defining cancers; SUD: Substance use disorder; HTN: Hypertension; MST: Metastatic solid tumor; DM: Diabetes mellitus; IA: Income assistance; AGE LINEAR: (age per 10-year increase - median age per 10-year increase in training set); CD4 LINEAR: (CD4 nadir per 100 cells - median CD4 nadir per 100 cells in training set); CD4 QUADRATIC: (CD4 nadir per 100 cells - median CD4 nadir per 100 cells in training set)²; YEARS ON ART LINEAR: (years on ART - median years on ART in training set); YEARS ON ART QUADRATIC: (years on ART - median years on ART in training set)²; YEARS ON ART CUBIC: (years on ART - median years on ART in training set)³

Step 2: We multiplied and summed each estimated regression coefficient from the mortality risk prediction index by the value of the corresponding covariates for each patient to get a risk coefficient.

Step 3:

Mortality risk prediction index scores were calculated using the formula below;

$$\text{Index score} = \frac{\text{Risk coefficient} - \text{minimum risk score where mortality is 0}}{\text{Maximum risk score where mortality is 100} - \text{minimum risk score where mortality is 0}} \times 100$$

Where:

Risk coefficient: Sum of products of coefficients and covariates

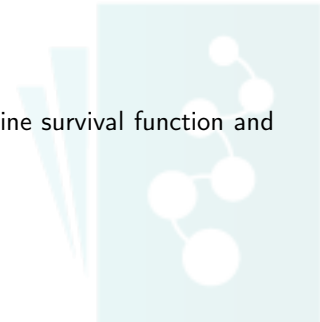
Minimum score: Mathematically derived score when mortality is 0

Maximum score: Mathematically derived score when mortality is 100

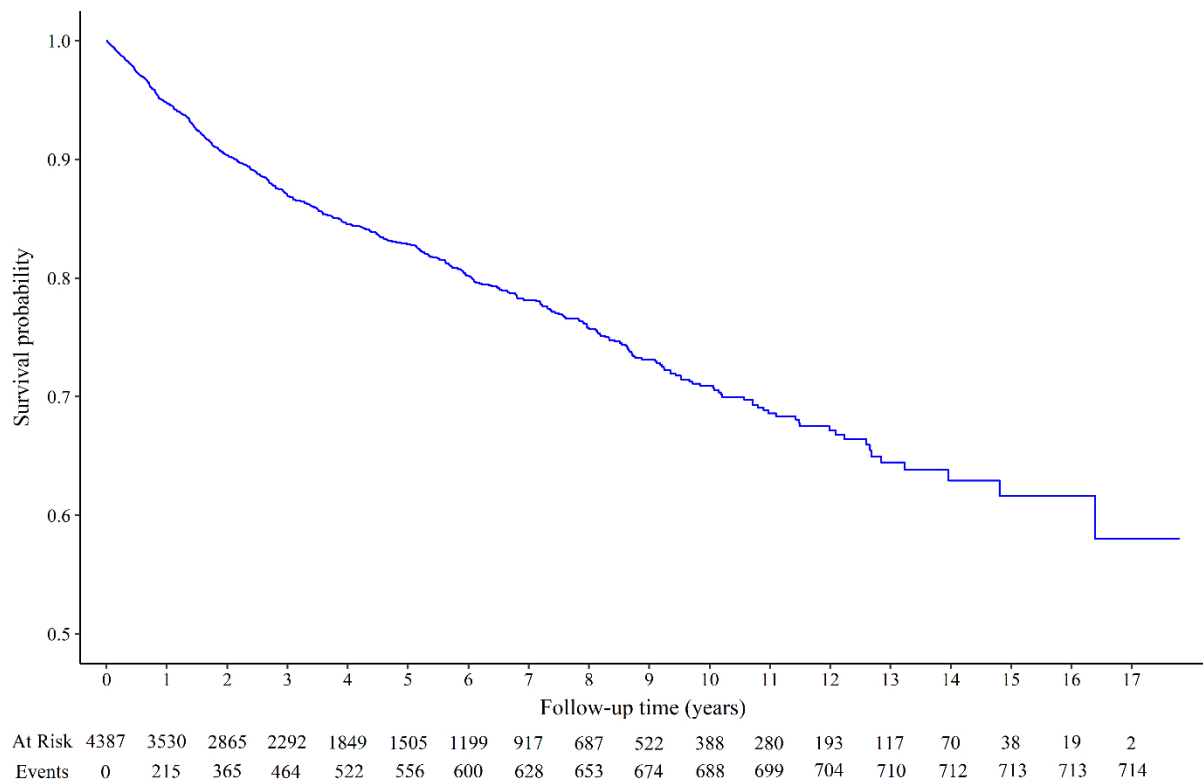
The mortality risk prediction index score obtained is between 0 and 100.

Step 4:

We used the Cox survival function to predict the one-year all-cause mortality probability using the baseline survival function and the sum of risk coefficients [28].



Supplementary Figure S2: Kaplan-Meier survival curve displaying survival probability from the index date to the study end period for people living with HIV in British Columbia, Canada, from 2002 to 2019



Supplementary Table S3: Cox proportional hazard model showing unadjusted hazard ratios and based on the training set

Variables	Unadjusted Hazard Ratio (95% Confidence Interval)
Cardiovascular disease (Ref: No)	
Yes	3.09 (2.46–3.88)
Chronic kidney disease (Ref: No)	
Yes	3.36 (2.64–4.28)
Diabetes mellitus (Ref: No)	
Yes	1.71 (1.29–2.26)
Chronic liver disease (Ref: No)	
Yes	3.34 (2.77–4.02)
Non-AIDS-defining cancers (Ref: No)	
Yes	2.66 (2.09–3.37)
Chronic obstructive pulmonary disease (Ref: No)	
Yes	3.70 (2.94–4.65)
Hypertension (Ref: No)	
Yes	1.51 (1.20–1.91)
Substance use disorder (Ref: No)	
Yes	3.72 (3.08–4.50)
Alzheimer's/Dementia (Ref: No)	
Yes	4.42 (3.26–5.99)
Personality disorder (Ref: No)	
Yes	1.95 (1.53–2.49)
Schizophrenia (Ref: No)	
Yes	1.72 (1.27–2.33)
Mood/Anxiety disorder (Ref: No)	
Yes	1.57 (1.31–1.89)
Rheumatoid arthritis (Ref: No)	
Yes	1.80 (0.74–4.34)
Metastatic solid tumor (Ref: No)	
Yes	4.57 (3.12–6.69)
Osteoporosis (Ref: No)	
Yes	2.59 (1.79–3.75)
Asthma (Ref: No)	
Yes	1.59 (1.28–1.98)
Peptic ulcer (Ref: No)	
Yes	1.66 (1.24–2.22)
Osteoarthritis (Ref: No)	
Yes	1.57 (1.13–2.16)
Plasma viral load suppression (Ref: Yes)	
No	2.77 (2.32–3.32)
Sex at birth (Ref: Male)	
Female	1.82 (1.50–2.21)
Income assistance (Ref: No)	
Yes	2.55 (2.12–3.05)
^a Age per 10-year increase	1.27 (1.18–1.38)
^b CD4 nadir per 100 cells/mm ³	0.70 (0.67–0.74)
^c Years on ART	0.94 (0.91–0.97)

Note: ART: Antiretroviral Therapy; ^aLinear form of age centered at the median (47 years); ^bPolynomial forms of CD4 nadir centered at the median (430 cells/mm³); ^cLinear form of years on ART centered at the median (4 years).

Supplementary Table S4: Cox proportional hazard model showing adjusted hazard ratios and Harrell's c-statistic based on full cohort

Characteristic	Adjusted Hazard Ratio (95% Confidence Interval)
Sex at birth (Ref: Male)	
Female	1.14 (0.96–1.36)
Cardiovascular disease (Ref: No)	
Yes	1.87 (1.51–2.31)
Chronic liver disease (Ref: No)	
Yes	1.48 (1.25–1.76)
Chronic kidney disease (Ref: No)	
Yes	1.62 (1.30–2.00)
Chronic obstructive pulmonary disease (Ref: No)	
Yes	2.04 (1.65–2.51)
Non-AIDS-defining cancers (Ref: No)	
Yes	1.79 (1.42–2.25)
Substance use disorder (Ref: No)	
Yes	2.52 (2.08–3.05)
Diabetes (Ref: No)	
Yes	1.17 (0.91–1.51)
Hypertension (Ref: No)	
Yes	0.86 (0.69–1.08)
Metastatic solid tumor (Ref: No)	
Yes	2.44 (1.73–3.45)
Plasma viral load suppression (Ref: Yes)	
No	1.60 (1.34–1.92)
Income assistance (Ref: No)	
Yes	1.28 (1.08–1.51)
^a Age per 10-year increase	1.35 (1.25–1.47)
^b CD4 nadir per 100 cells/mm ³	
Linear	0.82 (0.81–0.88)
Quadratic	1.02 (1.01–1.02)
^c Years on ART	
Linear	0.93 (0.89–0.97)
Quadratic	1.01 (0.99–1.02)
Cubic	1.00 (0.99–1.00)
†Unadjusted Harrell's c-statistic (95% CI)	0.8227 (0.8077–0.8377)
Optimism-adjusted Harrell's c-statistic (95% CI)	0.8188 (0.8180–0.8196)

Note: ART: Antiretroviral Therapy; ^aLinear form of age centered at the median (47 years); ^bPolynomial forms of CD4 nadir centered at the median (430 cells/mm³); ^cLinear form of years on ART centered at the median (4 years); †C-statistic derived from fitting the full data to the model.



Supplementary Table S5: Data table for predicted one-year all-cause mortality probabilities using our final Cox proportional hazard model vs observed Kaplan-Meier estimates of the one-year all-cause mortality probabilities by subgroups of mean mortality risk prediction index scores using the test dataset

Subgroups	Mean MRP _i score	n	Observed deaths	Observed mortality (95% CI)	Predicted mortality
1	33.2583	31	0	0.0000 (0.0000–0.0000)	0.0029
2	38.0982	171	0	0.0000 (0.0000–0.0000)	0.0053
3	42.5883	267	1	0.0045 (0.0006–0.0317)	0.0092
4	47.3144	278	7	0.0273 (0.0131–0.0565)	0.0164
5	52.4193	256	9	0.0376 (0.0197–0.0712)	0.0305
6	57.3821	171	6	0.0361 (0.0164–0.0787)	0.0554
7	62.5093	150	16	0.1107 (0.0693–0.1745)	0.1015
8	67.2987	72	10	0.1512 (0.0843–0.2632)	0.1755
9	72.2532	41	12	0.2927 (0.1781–0.4575)	0.2988
10	79.2786	26	9	0.3510 (0.2000–0.5671)	0.5693

Note: MRP_i: mortality risk prediction index; n: number of participants in the subgroups, CI: confidence interval.

Supplementary Table S6: Data table for predicted estimates of one-year survival probabilities vs observed adjusted and unadjusted Kaplan-Meier estimates of one-year survival probabilities by subgroups using the full dataset

Subgroups	Mean predictive survival	Unadjusted observed survival	95% CI of unadjusted observed survival	Adjusted observed survival
1	0.7587	0.7347	0.6939–0.7780	0.7400
2	0.9072	0.9018	0.8736–0.9308	0.9015
3	0.9432	0.9293	0.9048–0.9545	0.9282
4	0.9618	0.9830	0.9705–0.9956	0.9823
5	0.9725	0.9703	0.9539–0.9870	0.9697
6	0.9798	0.9832	0.9709–0.9956	0.9830
7	0.9851	0.9899	0.9800–0.9998	0.9889
8	0.9888	0.9973	0.9919–1.0000	0.9972
9	0.9918	1.0000	1.0000–1.0000	1.0000
10	0.9946	0.9971	0.9915–1.0000	0.9972

Note: CI: confidence interval.



Supplementary Table S7: Comparison of included and excluded participants from the study of people living with HIV in British Columbia, Canada, from 2002 to 2019

Characteristic	Included <i>n</i> = 4387	Excluded <i>n</i> = 1672	p-value*
Sex at birth, <i>n</i> (%)			
Male	3570 (81)	1336 (80)	0.2046
Female	817 (19)	336 (20)	
Death during study period, <i>n</i> (%)			
No	3673 (84)	1441 (86)	0.0204
Yes	714 (16)	231 (14)	
Continuous covariates	Median (Q1-Q3)	Median (Q1-Q3)	
Index year	2015 (2011–2017)	2011 (2007–2015)	<0.0001
Age at index date (years)	47 (38–64)	38 (31–46)	<0.0001
Year of ART initiation	2009 (2006–2013)	2009 (2006–2013)	0.9701
Years on ART (years)	4 (2–6)	5 (3–8)	<0.0001
Follow-up time (years)	3.3 (1.4–6.3)	3.3 (1.3–6.3)	0.8140

Note: Q1-Q3: 25th–75th percentiles; ART: Antiretroviral therapy.

*Bivariable comparisons of categorical variables were done by the chi-square test; Bivariable comparisons of continuous variables were done by the Wilcoxon's rank sum test. We compared participants in the final analytical sample to the total number of participants who had less than one year of follow-up, less than five years of administrative data, missing CD4 nadir and viral load. Results are significant at a p-value of <0.05.



Supplementary Table S8: Descriptive characteristics of males included in training and test datasets for people living with HIV in British Columbia, Canada, from 2002 to 2019

Variables	Full cohort (n = 4387) n (%)	Test dataset (n = 1463) n (%)	Training dataset (n = 2924) n (%)	p-value
Death during the study period				
No	3057 (86)	1013 (86)	2044 (86)	0.7671
Yes	513 (14)	166 (14)	347 (14)	
Cardiovascular disease				
No	3275 (92)	1088 (92)	2187 (92)	0.4438
Yes	295 (8)	91 (8)	204 (8)	
Chronic kidney disease				
No	213 (6)	71 (6)	142 (6)	0.9813
Yes	213 (6)	71 (6)	142 (6)	
Diabetes mellitus				
No	3307 (93)	1098 (93)	2209 (92)	0.4656
Yes	263 (7)	81 (7)	182 (8)	
Chronic liver disease				
No	3078 (86)	1000 (85)	2078 (87)	0.0982
Yes	492 (14)	179 (15)	313 (13)	
Non-AIDS-defining cancers				
No	3252 (91)	1074 (91)	2178 (91)	1.0000
Yes	318 (9)	105 (9)	213 (9)	
Chronic obstructive pulmonary disease				
No	3354 (94)	1121 (95)	2233 (93)	0.0554
Yes	216 (6)	58 (5)	158 (7)	
Hypertension				
No	3103 (87)	1034 (88)	2069 (87)	0.3570
Yes	467 (13)	145 (12)	322 (14)	
Substance use disorder				
No	2387 (67)	783 (66)	1604 (67)	0.7160
Yes	1183 (33)	396 (34)	787 (33)	
Alzheimer's/Dementia				
No	3480 (98)	1155 (98)	2325 (97)	0.2358
Yes	90 (2)	24 (2)	66 (3)	
Personality disorder				
No	3285 (92)	1072 (91)	2213 (93)	0.1041
Yes	285 (8)	107 (9)	178 (7)	
Schizophrenia				
No	3373 (95)	1122 (95)	2251 (94)	0.2387
Yes	197 (5)	57 (5)	140 (6)	
Mood/Anxiety disorder				
No	1938 (54)	667 (57)	1271 (53)	0.0586
Yes	1632 (46)	512 (43)	1120 (47)	
Rheumatoid arthritis				
No	3550 (99)	1171 (99)	2379 (99)	0.6696
Yes	20 (1)	8 (1)	12 (1)	
Metastatic solid tumor				
No	3503 (98)	1157 (98)	2346 (98)	1.0000
Yes	67 (2)	22 (2)	45 (2)	

Continued

Supplementary Table S8: Continued

Variables	Full cohort (n = 4387) n (%)	Test dataset (n = 1463) n (%)	Training dataset (n = 2924) n (%)	p-value
Osteoporosis				
No	3495 (98)	1158 (98)	2337 (98)	0.4173
Yes	75 (2)	21 (2)	54 (2)	
Asthma				
No	3182 (89)	1027 (87)	2155 (90)	0.0076
Yes	388 (11)	152 (13)	236 (10)	
Peptic ulcer				
No	3355 (94)	1099 (93)	2256 (94)	0.2038
Yes	215 (6)	80 (7)	135 (6)	
Osteoarthritis				
No	3383 (95)	1120 (95)	2263 (95)	0.7184
Yes	187 (5)	59 (5)	128 (5)	
CD4 nadir, cells/mm ³				
<50	121 (3)	43 (4)	78 (3)	0.3622
50-199	431 (12)	150 (13)	281 (12)	
200-349	725 (20)	221 (18)	504 (21)	
350+	2293 (64)	765 (65)	1528 (64)	
Plasma viral load suppression				
No	941 (26)	334 (28)	607 (25)	0.0663
Yes	2649 (74)	845 (72)	1784 (75)	
Income assistance				
No	2422 (68)	799 (68)	1623 (67)	0.9775
Yes	1148 (32)	380 (32)	768 (33)	
Continuous covariates	Median (Q1-Q3)	Median (Q1-Q3)	Median (Q1-Q3)	
Index year	2015 (2012-2018)	2015 (2011-2018)	2015 (2012-2018)	0.8813
Age at index date (years)	47 (40-55)	47 (39-54)	47 (40-55)	0.3884
Year of ART initiation	2009 (2006-2013)	2010 (2006-2013)	2009 (2006-2013)	0.2004
Years on ART at index date (years)	4.0 (2.0-6.0)	4.0 (2.0-6.0)	4.0 (2.0-70)	0.0999
CD4 nadir at index date (cells/mm ³)	440 (270-620)	430 (260-620)	440 (270-610)	0.8913
Follow-up time at end of follow-up (years)	3.2 (1.3-6.3)	3.3 (1.4-6.4)	3.2 (1.3-6.3)	0.4878

Legend: Q1-Q3: 25th - 75th percentiles; ART: antiretroviral therapy; p-value significance is set at 0.05.



Supplementary Table S9: Descriptive characteristics of females included in training and test datasets for people living with HIV in British Columbia, Canada, from 2002 to 2019

Variables	Full cohort (n = 817) n (%)	Test dataset (n = 284) n (%)	Training dataset (n = 533) n (%)	p-value
Death during the study period				
No	616 (75)	225 (79)	391 (73)	0.0769
Yes	201 (25)	59 (21)	142 (27)	
Cardiovascular disease				
No	757 (93)	267 (94)	490 (92)	0.3445
Yes	60 (7)	17 (6)	43 (8)	
Chronic kidney disease				
No	738 (94)	254 (89)	484 (91)	0.6123
Yes	79 (6)	30 (11)	49 (9)	
Diabetes mellitus				
No	757 (93)	264 (93)	493 (92)	0.9200
Yes	60 (7)	20 (7)	40 (8)	
Chronic liver disease				
No	574 (70)	196 (69)	378 (71)	0.6263
Yes	243 (30)	88 (31)	155 (29)	
Non-AIDS-defining cancers				
No	758 (93)	261 (92)	497 (93)	0.5720
Yes	59 (7)	23 (8)	36 (7)	
Chronic obstructive pulmonary disease				
No	744 (91)	259 (91)	485 (91)	1.0000
Yes	73 (9)	25 (9)	48 (9)	
Hypertension				
No	723 (89)	256 (90)	467 (88)	0.3364
Yes	94 (12)	28 (10)	66 (12)	
Substance use disorder				
No	280 (34)	80 (28)	200 (38)	0.0092
Yes	537 (66)	204 (72)	333 (62)	
Alzheimer's/Dementia				
No	792 (97)	277 (98)	515 (97)	0.6116
Yes	25 (3)	277 (98)	515 (97)	
Personality disorder				
No	693 (85)	245 (86)	448 (84)	0.4605
Yes	124 (15)	39 (14)	85 (16)	
Schizophrenia				
No	751 (92)	255 (90)	496 (93)	0.1341
Yes	66 (8)	29 (10)	37 (7)	
Mood/Anxiety disorder				
No	295 (36)	100 (35)	195 (37)	0.7543
Yes	522 (64)	184 (65)	338 (63)	
Rheumatoid arthritis				
No	799 (98)	277 (98)	522 (98)	0.9032
Yes	18 (2)	7 (2)	11 (2)	
Metastatic solid tumor				
No	805 (99)	279 (98)	526 (99)	0.8409
Yes	12 (1)	5 (2)	7 (1)	

Continued

Supplementary Table S9: Continued

Variables	Full cohort (n = 817) n (%)	Test dataset (n = 284) n (%)	Training dataset (n = 533) n (%)	p-value
Osteoporosis				
No	763 (93)	268 (94)	495 (93)	0.5019
Yes	54 (7)	16 (6)	38 (7)	
Asthma				
No	609 (75)	209 (74)	400 (75)	0.7111
Yes	208 (25)	75 (26)	133 (25)	
Peptic ulcer				
No	724 (89)	259 (91)	465 (87)	0.1142
Yes	93 (11)	25 (9)	68 (13)	
Osteoarthritis				
No	741 (91)	250 (88)	491 (92)	0.0733
Yes	76 (9)	34 (12)	42 (8)	
CD4 nadir, cells/mm ³				
<50	51 (6)	18 (6)	33 (6)	0.2698
50-199	159 (20)	52 (18)	107 (20)	
200-349	170 (20)	50 (18)	120 (23)	
350+	437 (54)	164 (58)	273 (51)	
Plasma viral load suppression				
No	330 (40)	112 (39)	218 (41)	0.7405
Yes	487 (60)	172 (61)	315 (59)	
Income assistance				
No	377 (46)	133 (47)	244 (46)	0.8308
Yes	440 (54)	151 (53)	289 (54)	
Continuous covariates	Median (Q1-Q3)	Median (Q1-Q3)	Median (Q1-Q3)	
Index year	2014 (2010-2017)	2015 (2011-2017)	2014 (2010-2017)	0.3925
Age at index date (years)	43 (36-52)	43 (36-50)	44 (36-52)	0.5497
Year of ART initiation	2009 (2005-2012)	2009 (2006-2012)	2009 (2005-2012)	0.7425
Years on ART at index date (years)	4.0 (2.0-6.0)	4.0 (2.0-6.3)	3.0 (2.0-6.0)	0.1320
CD4 nadir at index date (cells/mm ³)	370 (190-570)	410 (200-600)	360 (190-540)	0.1107
Follow-up time at end of follow-up (years)	3.4 (1.4-6.5)	3.4 (1.6-6.6)	3.3 (1.3-6.4)	0.3437

Legend: Q1-Q3: 25th–75th percentiles; ART: antiretroviral therapy; p-value significance is set at 0.05.



Supplementary Table S10: Cox proportional models for males showing adjusted hazard ratios based on the training dataset and Harrell's c-statistics for the training and test datasets

Variables	Model 1 aHR (95% CI)	Model 2 aHR (95% CI)	Model 3 aHR (95% CI)
Cardiovascular disease (Ref: No)			
Yes	2.01 (1.52-2.67)	1.74 (1.30-2.32)	1.80 (1.32-2.46)
Chronic liver disease (Ref: No)			
Yes	2.11 (1.65-2.71)	2.06 (1.61-2.63)	1.84 (1.43-2.37)
Chronic kidney disease (Ref: No)			
Yes	1.73 (1.27-2.36)	1.75 (1.29-2.38)	1.68 (1.22-2.31)
Chronic obstructive pulmonary disease (Ref: No)			
Yes	2.02 (1.49-2.75)	1.67 (1.24-2.24)	1.76 (1.30-2.39)
Non-AIDS-defining cancers (Ref: No)			
Yes	2.16 (1.59-2.93)	1.97 (1.44-2.69)	2.08 (1.51-2.87)
Substance use disorder (Ref: No)			
Yes	2.87 (2.27-3.62)	3.06 (2.41-3.87)	2.73 (2.11-3.54)
Alzheimer's/Dementia (Ref: No)			
Yes	2.24 (1.56-3.21)	2.26 (1.58-3.23)	1.67 (1.15-2.44)
Metastatic solid tumor (Ref: No)			
Yes	2.58 (1.61-4.15)	2.62 (1.64-4.19)	2.24 (1.38-3.63)
Diabetes mellitus (Ref: No)			
Yes	1.32 (0.94-1.84)		1.34 (0.94-1.92)
Rheumatoid arthritis (Ref: No)			
Yes	2.90 (0.92-9.12)	2.80 (0.89-8.80)	3.73 (1.17-11.94)
Asthma (Ref: No)			
Yes	0.79 (0.57-1.10)		
Hypertension (Ref: No)			
Yes			0.76 (0.55-1.06)
Income assistance (Ref: No)			
Yes			1.54 (1.20-1.96)
^a Age per 10-year increase		1.23 (1.10-1.38)	1.40 (1.24-1.58)
Plasma viral load suppression (Ref: Yes)			
No			1.41 (1.09-1.83)
^b CD4 nadir per 100 cells/mm ³			
Linear			0.85 (0.80-0.90)
Quadratic			1.02 (1.01-1.03)
^c Years on ART			
Linear			0.91 (0.86-0.97)
Quadratic			1.03 (1.01-1.05)
Cubic			1.00 (0.99-1.00)
Dataset statistic	Harrell's c- statistic (95% CI)	Harrell's c- statistic (95% CI)	Harrell's c- (95% CI)
Training dataset	0.7977 (0.7752-0.8202)	0.8002 (0.7771-0.8233)	0.8507 (0.8317-0.8697)
Test dataset	0.7598 (0.7221-0.7975)	0.7677 (0.7304-0.8050)	0.8069 (0.7734-0.8404)

Legend: ART: Antiretroviral Therapy; aHR: adjusted hazard ratio; CI: Confidence Interval; ^aLinear form of age centered at the median (47 years); ^bPolynomial forms of CD4 nadir centered at the median (440 cells/mm³); ^cLinear form of years on ART centered at the median (4 years).

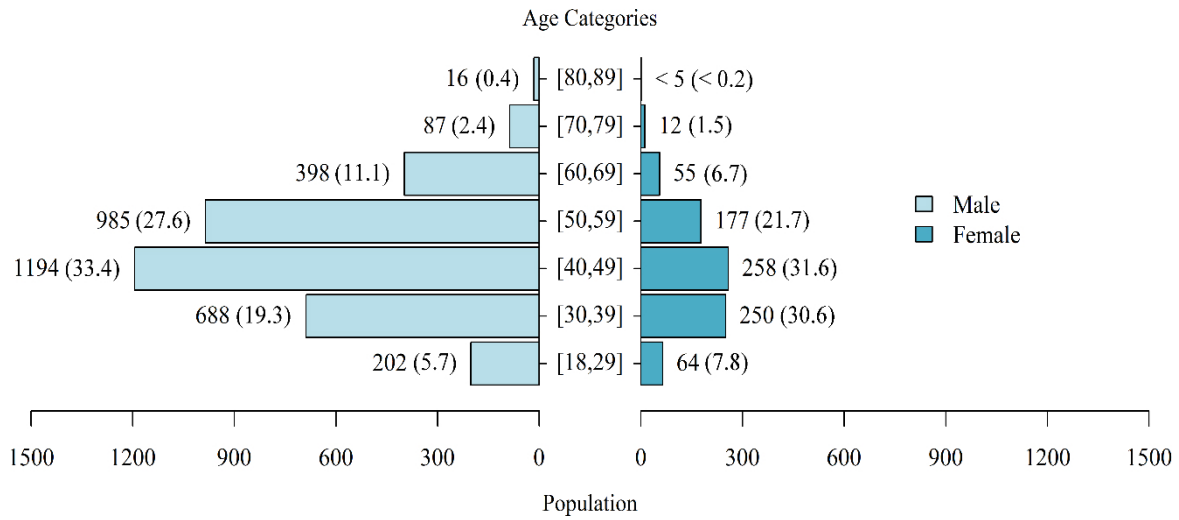
Supplementary Table S11: Cox proportional models for females showing adjusted hazard ratios based on the training dataset and Harrell's c-statistics for the training and test datasets

Variables	Model 1 aHR (95% CI)	Model 2 aHR (95% CI)	Model 3 aHR (95% CI)
Cardiovascular disease (Ref: No)			
Yes	2.11 (1.26-3.51)	2.10 (1.26-3.50)	1.83 (1.10-3.04)
Chronic kidney disease (Ref: No)			
Yes	2.52 (1.58-4.02)	2.41 (1.50-3.86)	1.73 (1.09-2.76)
Chronic obstructive pulmonary disease (Ref: No)			
Yes	2.43 (1.52-3.89)	2.20 (1.34-3.61)	2.96 (1.82-4.82)
Non-AIDS-defining cancers (Ref: No)			
Yes	2.07 (1.12-3.81)	1.99 (1.08-3.68)	2.37 (1.32-4.24)
Substance use disorder (Ref: No)			
Yes	2.60 (1.67-4.06)	2.76 (1.75-4.34)	2.48 (1.54-3.99)
Metastatic solid tumor (Ref: No)			
Yes	2.79 (0.91-8.60)	2.76 (0.88-8.60)	
Mood/Anxiety disorder (Ref: No)			
Yes	0.74 (0.52-1.07)	0.75 (0.52-1.08)	
Peptic ulcer (Ref: No)			
Yes	1.48 (0.94-2.34)	1.49 (0.94-2.36)	1.83 (1.18-2.83)
Hypertension (Ref: No)			
Yes	0.67 (0.38-1.18)	0.64 (0.36-1.13)	
^a Age per 10-year increase		1.11 (0.93-1.32)	1.24 (1.04-1.48)
Plasma viral load suppression (Ref: Yes)			
No			2.01 (1.35-3.00)
^b CD4 nadir per 100 cells/mm ³			0.80 (0.73-0.89)
^c Years on ART			0.96 (0.90-1.02)
Dataset	Harrell's c-statistic (95% CI)	Harrell's c-statistic (95% CI)	Harrell's c-statistic (95% CI)
Training dataset	0.7148 (0.6730-0.7566)	0.7198 (0.6772-0.7624)	0.7835 (0.7470-0.8200)
Test dataset	0.6978 (0.6278-0.7678)	0.6916 (0.6189-0.7643)	0.7492 (0.6777-0.8207)

Legend: ART: Antiretroviral Therapy; aHR: adjusted hazard ratio; CI: Confidence Interval; ^aLinear form of age centered at the median (44 years); ^bPolynomial forms of CD4 nadir centered at the median (360 cells/mm³); ^cLinear form of years on ART centered at the median (3 years).

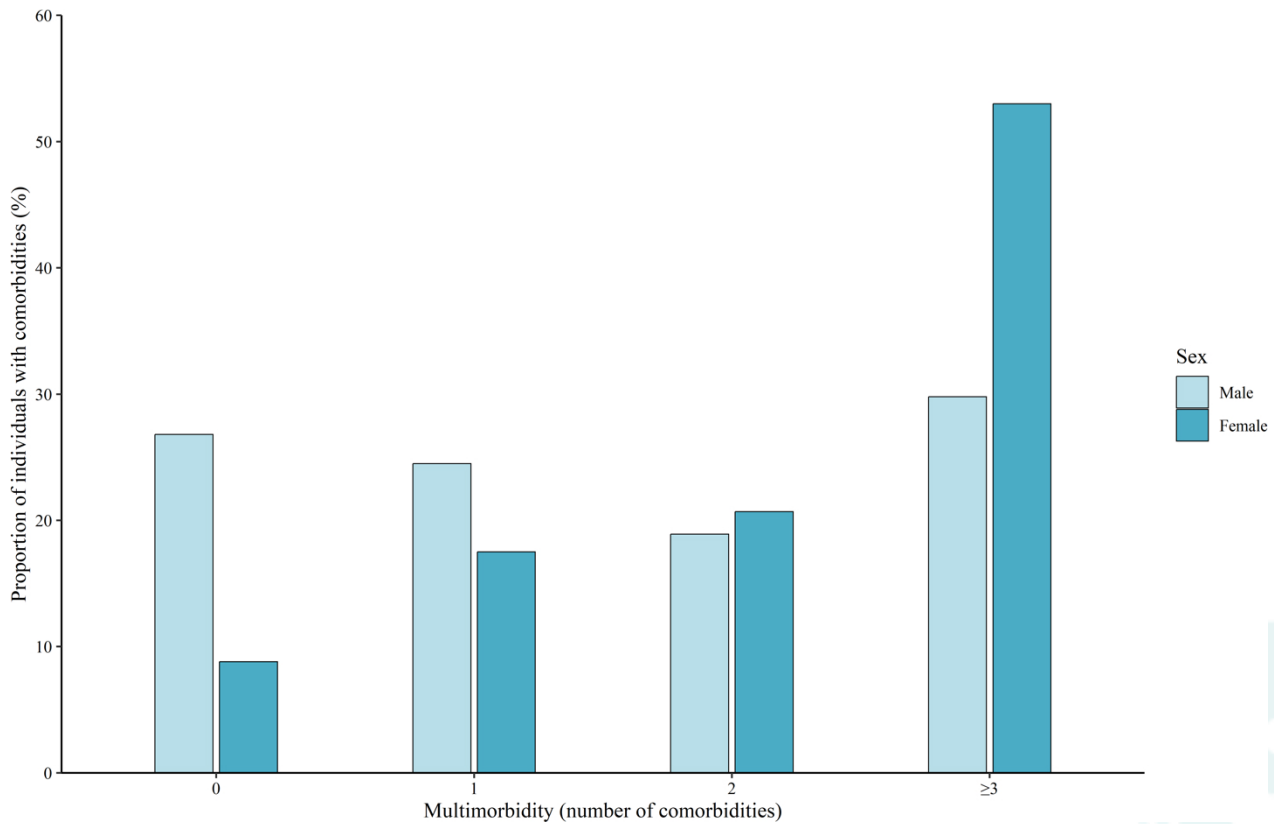


Supplementary Figure S3: Distribution of the full cohort (n = 4387) by age category and sex at birth of people living with HIV in British Columbia, Canada, from 2002 to 2019

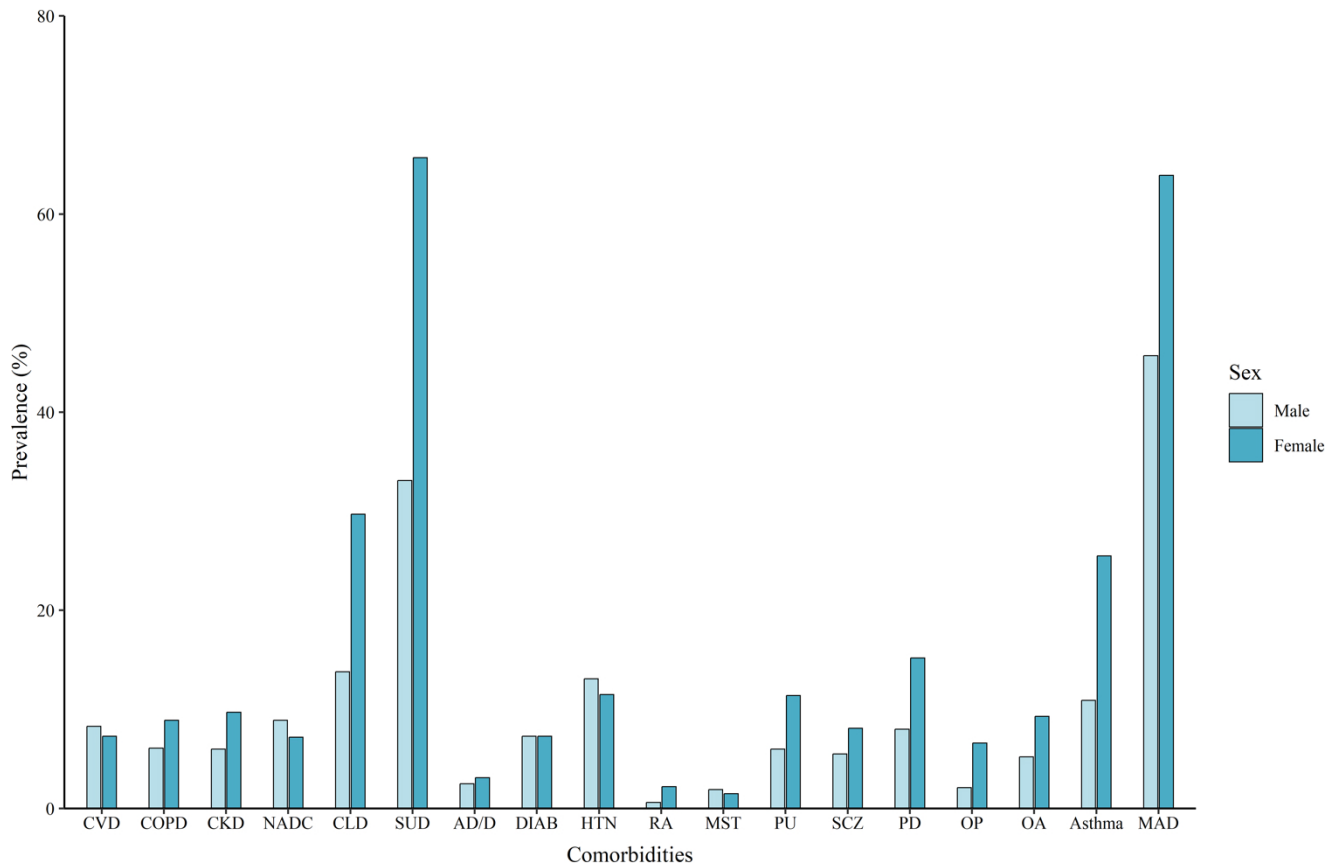


Note: Distribution expressed as number (%).

Supplementary Figure S4: The proportion of participants with or without comorbidities at the index date stratified by multimorbidity groups and by sex at birth among people living with HIV in British Columbia, Canada, from 2002 to 2019

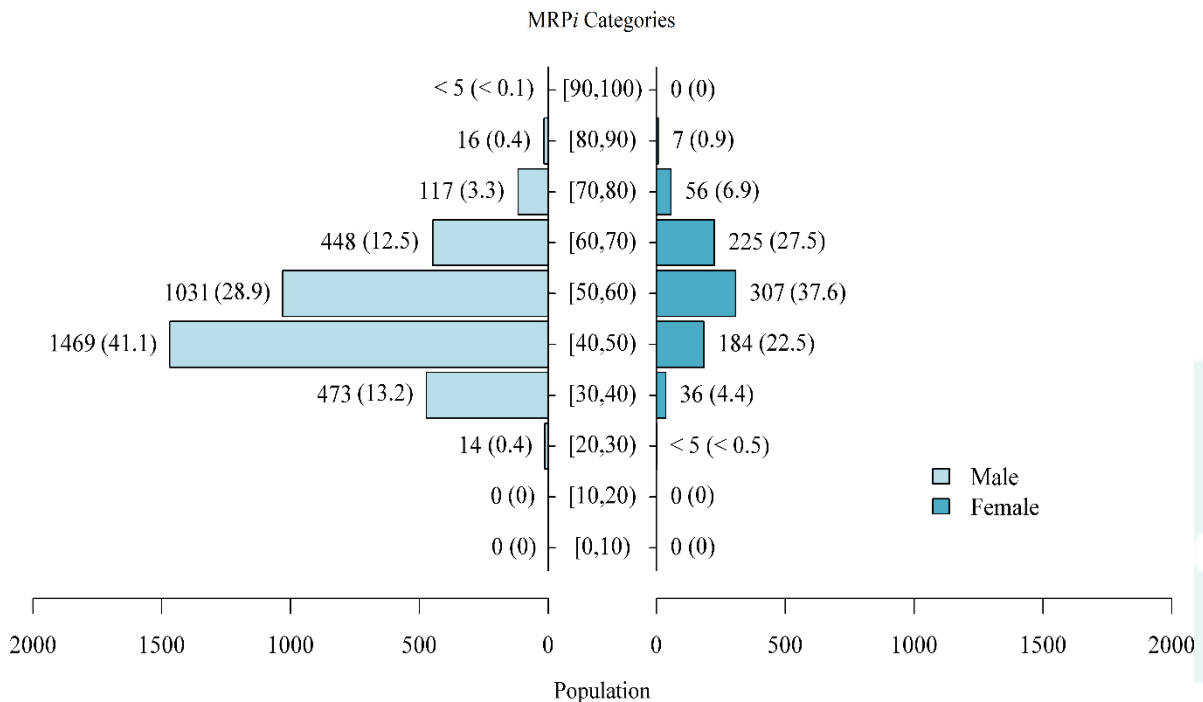


Supplementary Figure S5: Prevalence of comorbidities at the index date stratified by sex at birth among people living with HIV in British Columbia, Canada from 2002 to 2019



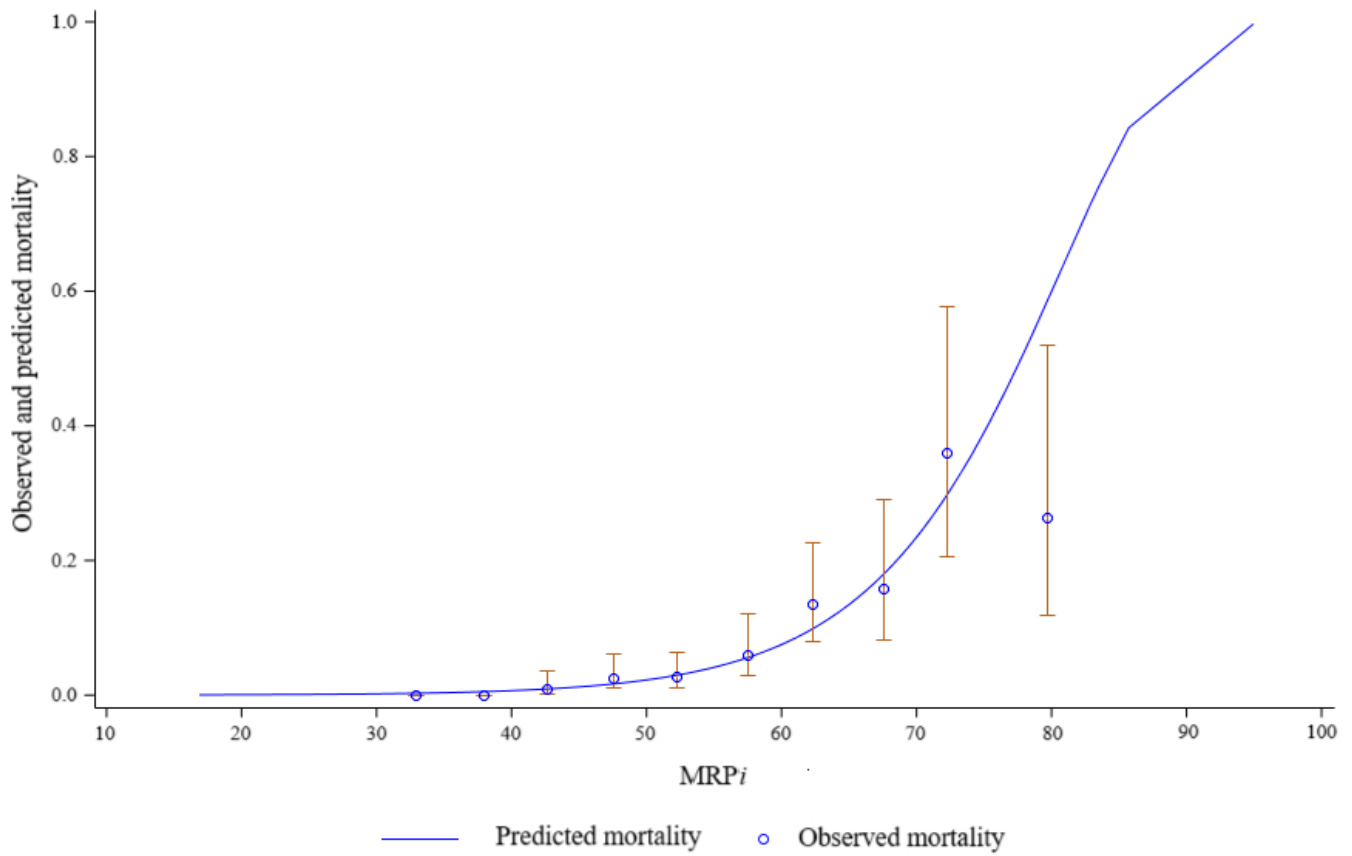
Note: CVD: Cardiovascular disease; COPD: Chronic obstructive pulmonary disease; CKD: Chronic kidney disease; NADC: Non-AIDS-defining cancers; CLD: Chronic liver disease; SUD: Substance use disorder; AD/D: Alzheimer's/Dementia; DIAB: Diabetes; HTN: Hypertension; RA: Rheumatoid arthritis; MST: Metastatic solid tumor; PU: Peptic ulcer; SCZ: Schizophrenia; PD: Personality disorder; OP: Osteoporosis; OA: Osteoarthritis; MAD: Mood/Anxiety disorder.

Supplementary Figure S6: Distribution of the full cohort (n=4387) by mortality risk prediction index score category and sex at birth of people living with HIV in British Columbia, Canada, from 2002 to 2019



Note: MRPi: Mortality risk prediction index. Distribution expressed as number (%).

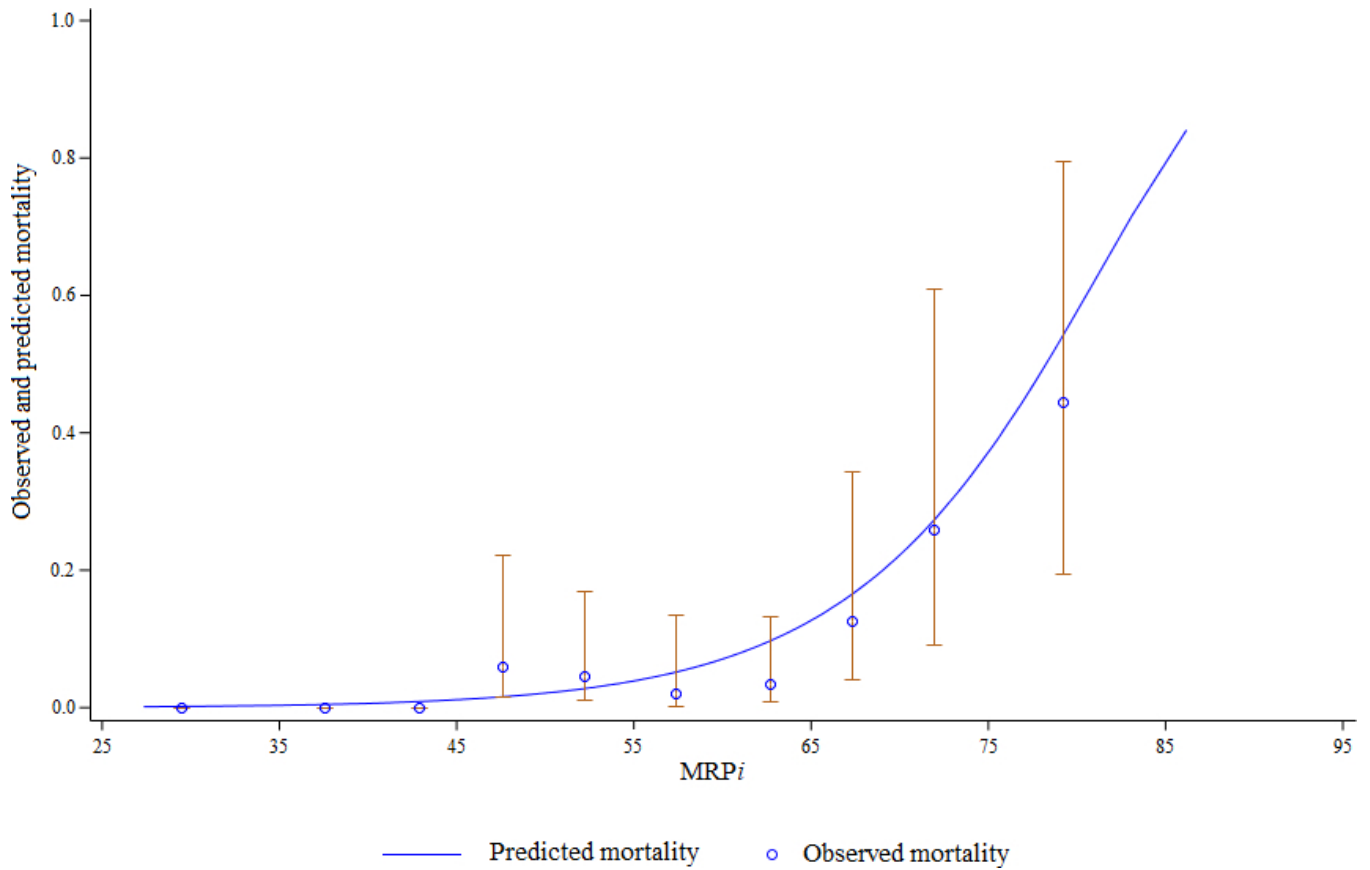
Supplementary Figure S7: Predicted one-year all-cause mortality for males using our final Cox model vs observed Kaplan-Meier estimates of the one-year all-cause mortality probability for males by subgroups of the one-year mortality risk prediction index (MRPi) using the test dataset



Legend: The observed one-year mortality probability is shown with 95% confidence intervals. Solid lines reflect predicted mortality probability calculated using the test dataset (n=1179). MRPi ranges from 0-100; x-axis begins from 10 because the minimum score in our sample is 17.



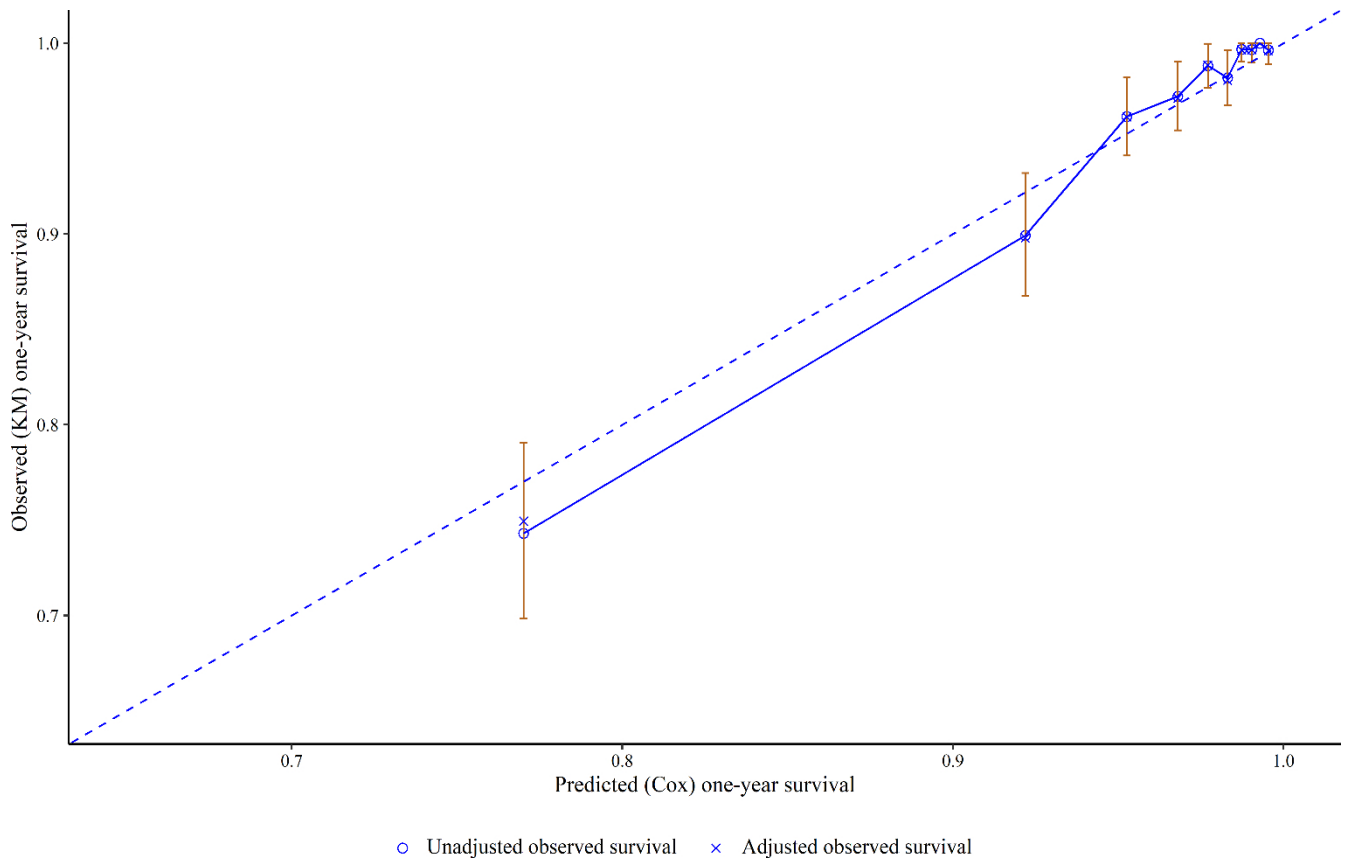
Supplementary Figure S8: Predicted one-year all-cause mortality for females using our final Cox model vs observed Kaplan-Meier estimates of the one-year all-cause mortality probability for females by subgroups of the one-year mortality risk prediction index (MRP_i) using the test dataset



Legend: The observed one-year mortality probability is shown with 95% confidence intervals. Solid lines reflect predicted mortality probability calculated using the test dataset (n=284). MRP_i ranges from 0-100; x-axis begins from 25 because the minimum score in our sample is 27.



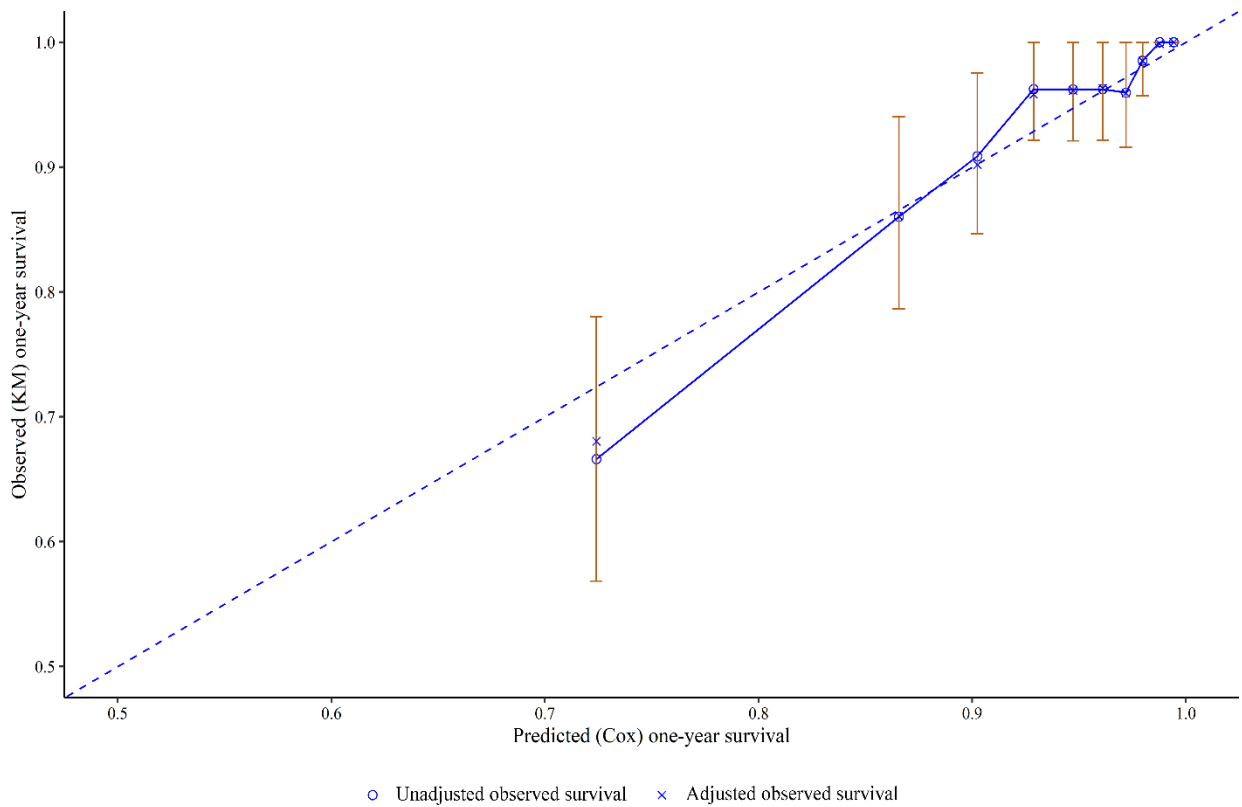
Supplementary Figure S9: Observed Kaplan-Meier estimates vs predicted estimates of one-year survival probabilities for males using the full dataset



Legend: KM; Kaplan-Meier. The observed one-year survival probability is shown with 95% confidence intervals. A solid line is drawn to join the two unadjusted observed survival points. The distance between a 45-degree (dotted) line with slope 1 and intercept 0 and the points reflects the difference between the observed and predicted survival probabilities. We did not have any observed survival probability <75%.



Supplementary Figure S10: Observed Kaplan-Meier estimates vs predicted estimates of one-year survival probabilities for females using the full dataset



Legend: KM; Kaplan-Meier. The observed one-year survival probability is shown with 95% confidence intervals. A solid line is drawn to join the two unadjusted observed survival points. The distance between a 45-degree (dotted) line with slope 1 and intercept 0 and the points reflects the difference between the observed and predicted survival probabilities. We did not have any observed survival probability <70%.



References

1. British Columbia Centre for Excellence in HIV/AIDS (BCCfE). Drug Treatment Program 2019 [Available from: <https://bccfe.ca/drug-treatment-program>].
2. Patterson S, Cescon A, Samji H, Cui Z, Yip B, Lepik KJ, et al. Cohort Profile: HAART Observational Medical Evaluation and Research (HOMER) cohort. *Int J Epidemiol*. 2015;44(1):58-67. <https://doi.org/10.1093/ije/dyu046>
3. BC Centre for Disease Control. [creator] 2017. Provincial HIV/AIDS surveillance database. BC Centre for Disease Control (BCCDC) [publisher]. Data extract. BCCDC (2017) [Available from: <http://www.bccdc.ca/our-services/service-areas/bccdc-public-health-laboratory>].
4. BC Ministry of Health. [creator] 2017. Medical Services Plan (MSP) Payment Information File. BC Ministry of Health [publisher]. Data extract. MOH (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
5. BC Ministry of Health. [creator] 2017. Pharmanet. BC Ministry of Health [publisher]. Data extract. MOH (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
6. BC Ministry of Health. [creator] 2017. Consolidation file (MSP registration & premium billing). BC Ministry of Health [publisher]. Data extract. MOH (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
7. Canadian Institute of Health Information. [creator] 2017. Discharge Abstract Database (Hospital Separations). BC Ministry of Health [publisher]. Data extract. MOH (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
8. BC Ministry of Health. [creator] 2017. Vital Statistics Deaths. BC Ministry of Health [publisher]. Data extract. BC Vital Statistics Agency (2017) [Available from: <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/data-access-health-data-central>].
9. 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). *Int J Epidemiol*. 2014;43(4):1073-81. <https://doi.org/10.1093/ije/dyu070>
10. Nosyk B, Colley G, Yip B, Chan K, Heath K, Lima VD, et al. Application and validation of case-finding algorithms for identifying individuals with human immunodeficiency virus from administrative data in British Columbia, Canada. *PLoS One*. 2013;8(1):e54416. <https://doi.org/10.1371/journal.pone.0054416>
11. Lapointe-Shaw L, Georgie F, Carlone D, Cerocchi O, Chung H, Dewit Y, et al. Identifying cirrhosis, decompensated cirrhosis and hepatocellular carcinoma in health administrative data: a validation study. *PLoS One*. 2018;13(8):e0201120. <https://doi.org/10.1371/journal.pone.0201120>
12. Roberts HW, Utuama OA, Kleven M, Teshale E, Hughes E, Jiles R. The contribution of viral hepatitis to the burden of chronic liver disease in the United States. *Am J Gastroenterol*. 2014;109(3):387-93. <https://doi.org/10.1038/ajg.2013.477>
13. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition - Alzheimer's/Dementia: British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
14. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition - Diabetes Mellitus: British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
15. Williamson T, Green ME, Birtwhistle R, Khan S, Garies S, Wong ST, et al. Validating the 8 CPCSSN case definitions for chronic disease surveillance in a primary care database of electronic health records. *Ann Fam Med*. 2014;12(4):367-72. <https://doi.org/10.1370/afm.1644>
16. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition - Hypertension: Ministry of Health British Columbia; 2015 [updated 2017. 2015: [Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].

17. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Ischemic Heart Disease: British Columbia Ministry of Health; 2015 [updated 2017. 2015: [Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
18. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Stroke, Hospitalized Transient Ischemic Attack (TIA) British Columbia Ministry of Health; 2015 [updated 2017. 2015: [Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
19. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Stroke, Hospitalized BC Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
20. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Heart Failure: BC Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
21. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Acute Myocardial Infarction (AMI), Hospitalized: BC Ministry of Health 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
22. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Chronic Kidney Disease: British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
23. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Dialysis: British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
24. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Kidney Transplant: British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
25. Chronic Disease Information Working Group. BC Chronic Disease and Selected Procedure Case Definitions, Case Definition – Chronic Obstructive Pulmonary Disease British Columbia Ministry of Health; 2015 [updated 2017. 2015:[Available from: <https://www2.gov.bc.ca/assets/gov/health/conducting-health-research/data-access/chronic-disease-registries-case-definitions.pdf>].
26. BC Cancer Statistics and Reports. Cancer Type and Case definitions BC Cancer Agency; 2021 [updated 2021. Available from: http://www.bccancer.bc.ca/statistics-and-reports-site/Documents/cancer%20definitions%20table_20210201.pdf].
27. Chronic Disease Information Working Group. British Columbia Chronic Disease Registries Case Definition - Substance Use Disorder: British Columbia Ministry of Health; 2018 [updated 2022. 2018:[Available from: <http://www.bccdc.ca/resource-gallery/Documents/Chronic-Disease-Dashboard/substance-use-disorder.pdf>].
28. Harrell Jr FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Second ed. New York (US): Springer; 2001.

