UCLA
UCLA Electronic Theses and Dissertations

Title

Refining Atherosclerotic Cardiovascular Risk Assessment Tool for Prediction of Subclinical Atherosclerosis, All-cause Mortality, and Non-HIV Mortality among People Living with HIV in the Multicenter AIDS Cohort Study and the Women's Interagency HIV S...

Permalink

Author

Tanasugarn, Lokachet TIME

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Refining Atherosclerotic Cardiovascular Risk Assessment Tool for Prediction of

Subclinical Atherosclerosis, All-cause Mortality, and Non-HIV Mortality

among People Living with HIV in the Multicenter AIDS Cohort Study and the Women's

Interagency HIV Study Combined Cohort Study

A dissertation submitted in partial satisfaction of the

requirement for the degree of Doctoral of Philosophy

in Epidemiology

by

Lokachet Tanasugarn

2021

ABSTRACT OF THE DISSERTATION

Refining Atherosclerotic Cardiovascular Risk Assessment Tool for Prediction of

Subclinical Atherosclerosis, All-cause Mortality, and Non-HIV Mortality

among People Living with HIV in the Multicenter AIDS Cohort Study and the Women's

Interagency HIV Study Combined Cohort Study

by

Lokachet Tanasugarn

Doctor of Philosophy in Epidemiology

University of California, Los Angeles, 2021

Professor Sung-Jae Lee, Chair

These studies aimed to build prediction models for cardiovascular-related outcomes among people living with HIV and compared model performance with previous standard models. The three outcomes were new subclinical atherosclerotic plaques, all-cause mortality, and non-HIV mortality. The two standard models for comparison were the pooled cohort equations and the Veteran Aging Cohort Study (VACS) Index. We used the MACS/WIHS Combined Cohort Study to build each model in the whole data and separately as sex-specific models. During the model building process, we incorporated machine learning

techniques, including the elastic net regularization and the boosting ensemble approach, for better prediction.

The results from the three studies point in the same direction that the new model offers better predictive performance than the previous benchmarking models. However, only the 10-year sex-specific models and the 10-year MWCCS model for all-cause mortality satisfied all the performance measures and would be recommended to further pursue external validation.

The dissertation of Lokachet Tanasugarn is approved.

Roger Detels

Matthew Budoff

Warren Scott Comulada

Sung-Jae Lee, Committee Chair

University of California, Los Angeles

2021

Table of Contents

List of Figures

List of Tables

Chapter 3: Manuscript #1

Chapter 4: Manuscript #2

Chapter 5: Manuscript #3

List of Acronyms

AIDS	Acute Immunodeficiency Syndrome
ASCVD	Atherosclerotic Cardiovascular Disease
AZT	Azidothymidine
CVD	Cardiovascular Disease
CDC	Center of Disease Control
cART	Combined Anti-Retroviral Therapy
CRP	C-reactive protein
D:A:D	Data Collection on Adverse Events of Anti-HIV Drugs
DBP	Diastolic Blood Pressure
DALY	Disability Adjusted Life Year
HDL-C	High Density Lipoprotein Cholesterol
HIV	Human Immunodeficiency Virus
IL-6	Interleukin-6
IMT	Intima Media Thickness
LDL-C	Low Density Lipoprotein Cholesterol
MSM	Men who have sex with men
MACS	Multicenter AIDS Cohort Study
MWCCS	Multicenter AIDS Cohort Study/Women's Interagency HIV Study Combined Cohort Study
MI	Myocardial Infarction
NHLBI	National Heart, Lung, and Blood Institute
NIAID	National Institute of Allergy and Infectious Diseases
NRI	Net reclassification index
NRI+	Net reclassification index of event
NRI-	Net reclassification index of non-event

NCD	Non-Communicable Disease
NNRTI	Non-Nucleoside/Nucleotide Reverse Transcriptase Inhibitors
NA-ACCORD	North American AIDS Cohort Collaboration on Research and Design
NRTI	Nucleoside/Nucleotide Reverse Transcriptase Inhibitors
PLHIV	People Living With HIV
PWID	People Who Inject Drugs
PCE	Pooled Cohort Equations
PROCAM	Prospective Cardiovascular Muenster Study
PI	Proteinase Inhibitor
RAMA-EGAT	Ramathibodi-Electricity Generating Authority of Thailand
sCD	Soluble Cluster of Difference
SCORE	Systematic Coronary Risk Evaluation
SBP	Systolic Blood Pressure
TC	Total Cholesterol
TG	Triglyceride
TNF	Tumor Necrosis Factor
US	United States
UCLA	University of California Los Angeles
VACS	Veteran Aging Cohort Study
WIHS	Women's Interagency HIV Study

Acknowledgments

I wish to thank the following people and institutions for the following reasons.

Wendy Aft	Her willingness to help Fogarty students with several issues both in the school and beyond
Joy Miller	Her swift assistant regarding any official administrative issue at the department
Colleagues at the Fielding School of Public Health	All the shared experience throughout this journey, notably the qualification exam study group
Thai UCLA Epidemiology Student Alumni	Ongoing support and networking even before the program of study
Powell Library	Excellent study space including Night Powell, Group study room "B," and Computer Lab
UCLA Recreation Centers	24/7 exercise at the John Wooden Center and a lot of lap swimming at the park pool
UCLA Vending Machines	Healthy snacks and (not-so-healthy) beverages

VITA

Education

MD from Mahidol University (First Class Honor)   2015

Awards and honors

- **Mimi and Roger Detels fellowship** from the Department of Epidemiology, FSPH, UCLA   2020 - 2021

- **Epidemiology student service award** from the Department of Epidemiology, FSPH, UCLA   2021

- **Student leadership award** from the Department of Epidemiology, FSPH, UCLA   2020

- **PHSA bouquet award** from the Public Health Student Association (PHSA), FSPH, UCLA   2019

- **The Fogarty AIDS International Training and Research Program** at the Department of Epidemiology, FSPH, UCLA   2017 - 2020

Research experience

- **Graduate Student Researcher**:   2020 - 2021
  - Refining Atherosclerotic Cardiovascular Risk Assessment Tool for Prediction of Subclinical Atherosclerosis, All-cause Mortality, and Non-HIV Mortality among PLHIV in the Multicenter AIDS Cohort Study and the Women's Interagency HIV Study Combined Cohort Study

Teaching assistant experience

- **EPIDEM 220:** Principles of Infectious Disease Epidemiology   Winter 2021

- **EPIDEM 200C:** Method III - Analysis   Spring 2019, Spring 2020

- **EPIDEM 200B:** Method II – Prediction and Validity   Winter 2019, Winter 2020

- **EPIDEM 200A:** Method I – Basic Concepts and Study Designs   Fall 2018

CHAPTER 1

INTRODUCTION

1. Introduction

1.1 Epidemiology of HIV

1.1.1 Global Epidemiology of HIV

Since the occurrence of a cluster of previously healthy MSM with *Pneumocystis jirovecii* in 1981, HIV and AIDS have been one of the most important infectious disease threats globally[1]. In 2002, the US suffered the economic burden of $36.4 billion from HIV/AIDS[2]. Economic studies from subsequent years also showed consistent higher cost among PLHIV compared to general population counterpart with more pronounced burden on comorbidities care[3,4]. From the start of the epidemic until the end of 2019, 75.7 million people have been infected with HIV, and 32.7 million people have died from AID-related illness worldwide[5]. Even in 2019 alone, there were 38 million PLHIV in which 1.7 million people were newly infected, and 690,000 people died from AIDS-related illness globally[5]. In terms of transmission risk, MSM, PWID, sex workers, and transgender people confers 26, 29, 30, and 13 times higher risk of HIV transmission, accordingly[5].

1.1.2 HIV Epidemiology in the US

In the US during 2014 - 2018, the overall number of HIV infections decreased from 40,836 to 37,968 diagnoses[6]. Despite this decrease, new HIV diagnosis affected specific subgroups differentially by race, region, and transmission category[6]. In 2018, 69% of HIV diagnoses were from Black/African Americans and Hispanic/Latino, who constitute only 31% of the US population[6]. Based on the regional difference, 51% of the new diagnosis in 2018 was in the Southern region of the US[6]. MSM and heterosexual contact accounted for 66% and 24% for the transmission category, making them the first and the second most common transmission

routes[6]. Within the same timeframe, deaths of PLHIV remained stable, with 16,619 deaths in 2014 and 15,821 deaths in 2018[6]. This downward trend of all-cause death could also be seen from a longer duration from 2010 to 2017[7]. The steady decline resulted from decreased HIV-related death in the more recent years[7]. At the end of 2018, a total of 1,040,352 PLHIV were in the US[6]. To meet the 2030 goal of ending the HIV epidemic, the US still needs to implement changes in several areas ranging from clinical practice to policy administration to target the remaining differential burden of HIV/AIDS[8,9].

1.1.3 Rise of CVD among PLHIV

The invention of AZT in 1986 and other antiretroviral therapies that were later incorporated into the cART was the real game-changer in the fight against HIV/AIDS in the early phase of the epidemic[1]. The highly effective combined regimen turned HIV/AIDS from an inescapable death sentence into a controllable long-term medical condition[1]. The cART has become the standard of care for HIV/AIDS in the US since 1996, leading to a longer life expectancy of PLHIV[1]. Consequently, this prolonged survival allows PLHIV to experience aging and non-communicable diseases[1]. This change of mortality pattern was illustrated from the HIV outpatient study from 1996 to 2004, where the percentage of death from non-AIDS-defining-illness proportionally increased as opposed to a decrease in AIDS-related death[10]. The non-AIDS-defining-illnesses in this study encompassed CVD, hepatic disease, pulmonary disease, and non-AIDS malignancies[10]. To specifically focus on CVD mortality among PLHIV, Feinstein et al. utilized the Wide-Ranging Online Data for Epidemiology Research (WONDER) from the CDC to look at the national level proportionate CVD mortality among PLHIV from 1999 to 2013[11]. The findings revealed that CVD mortality among the HIV-positive group escalated significantly while the overall mortality decreased from 1999 to 2013,

as shown in Figure 1[11]. This emerging trend emphasized the need to prepare for the upcoming

CVD burden among PLHIV where traditional treatment and prevention tools might not be

applicable.

Figure 1. Proportionate mortality for CVD of all deaths within the general population, inflammatory poly arthropathy population, and HIV-infected population (Feinstein et al. 2016) [11]
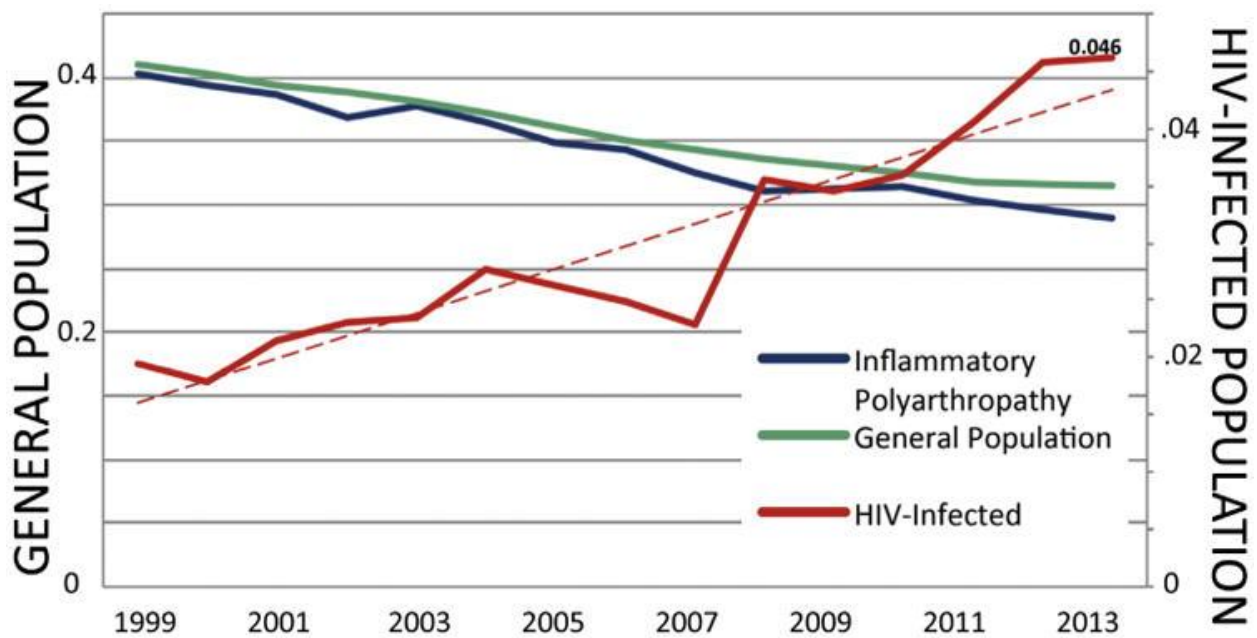
1.2 Atherosclerotic Cardiovascular Disease among PLHIV

1.2.1 CVD Terminology and Classification

CVDs are multifactorial and presented through a spectrum of clinical presentations in several organ systems[12]. Major CVDs are coronary heart disease, hypertension, heart failure, stroke, and peripheral artery disease[12]. Among PLHIV, the risk of developing MI, heart failure, pulmonary hypertension, and thrombosis is higher than in the general population[13]. Atherosclerotic Disease (ASCVD) is a group of CVDs associated with the formation of atherosclerotic plaque[13]. The plaque results from inflammation and lipid deposition in the vessel wall that eventually impede blood flow leading to ischemia of the related end organs downstream from the affected vessel. ASCVD outcome can be classified into hard and soft ASCVD outcome[13]. The hard ASCVD outcome encompasses non-fatal myocardial infarction, coronary heart disease death, non-fatal stroke, and fatal stroke[13]. Soft ASCVD endpoint includes angina, coronary disease requiring percutaneous intervention or bypass surgery, and subclinical measures[13]. These subclinical measures are carotid intimal medial thickness, coronary calcium, angiographic stenosis, and brachial ultrasound flow mediated dilation[13].

1.2.2 ASCVD and HIV

Globally, the burden of ASCVD among PLHIV has been gradually rising. A systematic review and meta-analysis by Shah et al. estimated that the relative risk of ASCVD among PLHIV increases by 50% to double the risk compared to people without HIV worldwide[14]. Additionally, the population-attributable fraction of ASCVD from HIV increased from 0.36% in 1990 to 0.92% in 2015[14]. Furthermore, DALYs from HIV-related ASCVD increased from 0.74 million in 1990 to 2.06 million in 2015, reflecting an approximately 3-fold increase in disease burden[14]. The time trend also persisted after stratification by sex[14]. This heightened

DALY exhibited geographical variation with most of the burden in Sub-Saharan Africa and Asia Pacific regions[14]. These estimated burdens mapped by country are presented in Figure 2—all of the estimations point towards the need for proper ASCVD care among PLHIV[14].

Figure 2. Global burden of atherosclerotic cardiovascular disease in people living with HIV(A) Population-attributable fraction by country and (B) disability-adjusted life-years per 100, 000 people by country (Shah et al.2018) [14]
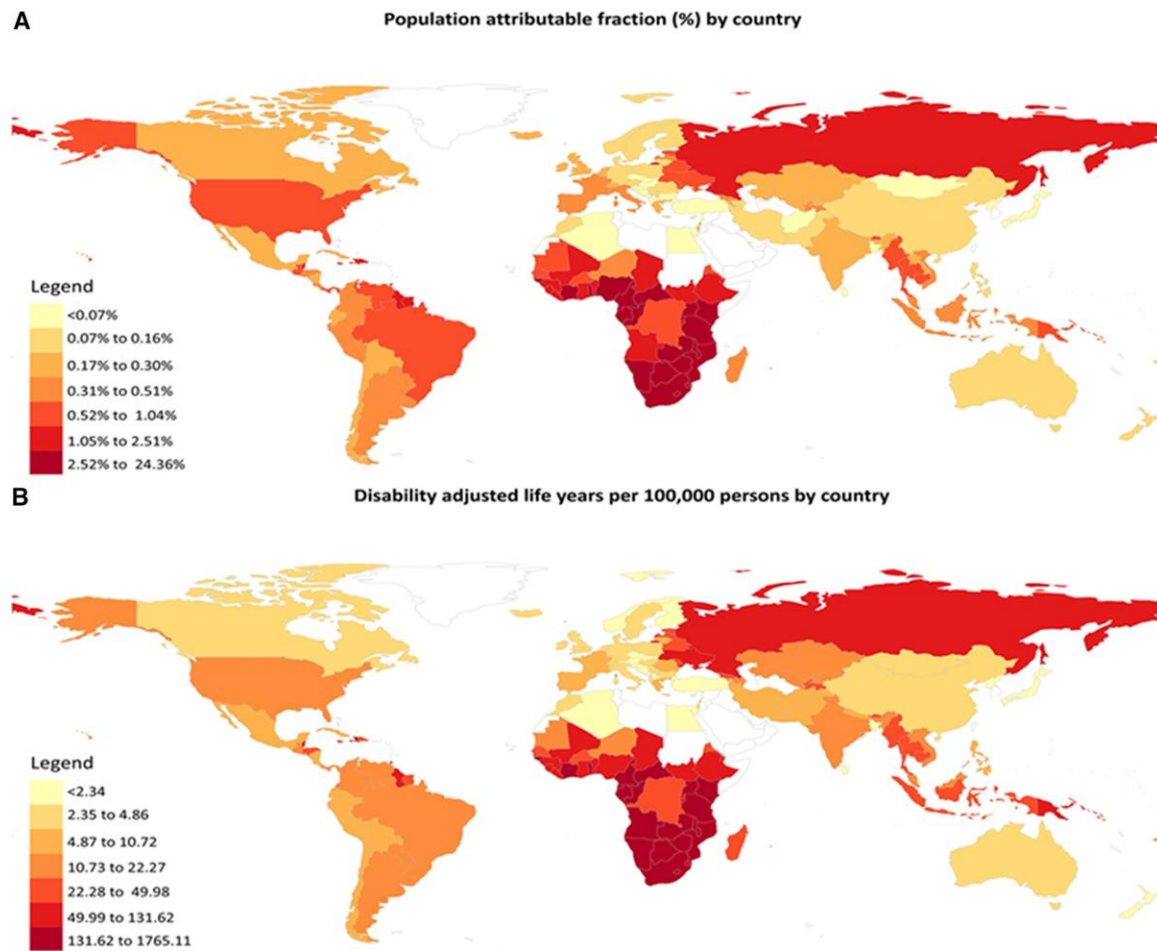
1.2.3 Factors associated with the Atherosclerotic Process among PLHIV

Despite advancements in exploring excess ASCVD risk among PLHIV, the exact explanation for this phenomenon is still unclear. The unique atherosclerotic process has been studied from several angles, including demographic profile, cART, inflammation biomarkers, and imaging findings.

1.2.3.1 Association between Demographic Factors and Atherosclerosis among PLHIV

Traditional ASCVD risk factors such as smoking, hypertension, diabetes mellitus, and lipid profile among PLHIV differ from the general population. PLHIV tend to develop ASCVD at a younger age than the general population. This early manifestation has been hypothesized as related to accelerated aging among PLHIV. For smoking, a representative PLHIV sample in the US had 42% current smokers compared to 20.6% in the general population showing a distinctively high prevalence of smoking[15]. Globally, PLHIV were estimated to have an overall prevalence of hypertension of 25.2%[16]. The prevalence differs by HIV treatment status with 34.7% for cART-treated and 12.7% for cART-naïve PLHIV[16]. For lipid profile, PLHIV often have a normal level of low-density lipoprotein cholesterol (LDL-C), the recent cART regimen does not appear to elevate lipid levels[13]. It has been suggested that atherogenic dyslipidemia in PLHIV relates to innate immune activation and other factors that need further investigation[13]. In terms of glucose dysregulation, diabetes mellitus also confers around 2.4 times the risk of ASCVD among PLHIV compared to the non-HIV population[17].

1.2.3.2 Association between cART use and Atherosclerosis among PLHIV

cART is an important factor for ASCVD events because uncontrolled HIV viremia is related to a higher chance of MI and because some antiretroviral medications are associated with

an ASCVD event[13]. Analysis of protease inhibitor (PI) use from the D:A:D cohort was related to additional MI risk after adjusting for cholesterol level[13]. Subsequent studies have confirmed the findings of an additional ASCVD risk, a class effect for PI except for atazanavir[13]. Among the more recent NRTIs Abacavir is associated with higher MI risk from the D:A:D cohort[18,19], the Kaiser Permanente California Health System study[20], and the NA-ACCORD study[21].

1.2.3.3 Association between Inflammatory Biomarkers and Atherosclerosis among PLHIV

Chronic inflammation from long-standing immune activation is the hallmark in the natural history of HIV & AIDS[1]. This chronic inflammation is also predictive of mortality, Non-AIDS events, and ASCVD[13]. A high level of several biomarkers, including IL-6, TNF α-1, and α-2, monocyte activation marker sCD163 and sCD14, is associated with the progression of atherosclerosis, and the level of these markers remains high among virally suppressed PLHIV[13].

1.2.3.3 Association between Imaging findings and Atherosclerosis among PLHIV

Imaging modalities, namely carotid ultrasound for measuring IMT and CT scan for evaluation of CAC score, allows evaluation of the subclinical atherosclerotic process[13]. PLHIV had thicker IMT in a cross-sectional study and a higher rate of IMT thickening from a longitudinal setting than the non-HIV group[13]. This differential progression is also confirmed based on the CAC score[22]. Among PLHIV, the atherosclerotic plaque tends to be noncalcified[23] with evidence of coronary artery remodeling[24]. These lesions possess a higher risk of plaque rupture, leading to a higher risk of an ASCVD event among PLHIV[13]. The difference in the distribution of traditional ASCVD risk factors, influence from cumulative

cART exposure, elevated soluble inflammatory markers, and the faster pace of subclinical progression result in the need for more refined ASCVD treatment and prevention strategies among PLHIV[13].

1.3 CVD risk assessment tool and their application in PLHIV

1.3.1 Concept of ASCVD Prevention

The key principle of primary ASCVD prevention is that the intensity of the prevention effort has to match the absolute risk of the patient[12]. Lifestyle modification should be recommended for every patient, while high-risk patients need more intensive lifestyle intervention and pharmacotherapy[12]. The continuously updated prediction ASCVD risk prediction models have been the cornerstone of preventive cardiology practice worldwide[12]. Several versions of a predictive model have been developed to meet this need in response to the broad clinical spectrum of CVD and heterogeneity in risk profile in different populations.

1.3.2 ASCVD Risk Prediction Model among General Population

Several ASCVD risk prediction models for the general population have been developed and validated. One of the most well-known risk prediction models is the PCE[25]. The PCE is a sex-race-specific 10-year ASCVD risk prediction model among the US general population. A working group of experts extensively planned to revise a new ASCVD prediction model in the US to be more representative and more rigorously evaluated. This plan led to data pooling across five community-based cohorts: the Framingham original cohort study, the Framingham offspring cohort study, the Atherosclerosis Risk in Communities Study, the Cardiovascular Health Study, and the Coronary Artery Risk Development in Young Adults study. The data pooling resulted in 11,240 White women (902 ASCVD events), 9,098 White

9

men (1,259 events), 2,641 African American women (290 events), and 1,647 African American men (238 events) in their model development data. The 10-year risk is the risk of non-fatal MI, coronary heart disease, non-fatal or fatal stroke among people free of ASCVD at the beginning of the specified period. The actual model is a proportional hazard model with age, treated or untreated systolic blood pressure, total cholesterol, high-density lipoprotein cholesterol, smoking status, and diabetes status as covariates. The working group tested the interaction of each predictor with age and retained the interaction terms based on p-values. Internal validation was done by cross-validation. External validation was carried out in the most recent examination cycles from ARIC and Framingham and two external cohorts: the Multi-Ethnic Study of Atherosclerosis (MESA) and the REasons for Geographic And Racial Differences in Stroke study (REGARDS). The working group launched the PCE in 2013 along with its risk calculator[25].

Notably, the PCE is a rigorously designed and evaluated prediction model among the general population. However, it lacks variables needed to capture related risks among PLHIV, and the values of each traditional predictor plausible differ from the general population. Their main statistical advantage was the enormous amount of data resulting from the pooling. Nevertheless, they could have split the initial data into training and testing data sets resembling the development and testing dataset for the internal validation process. The splitting helps address overfitting issue and reflect the real application of newly developed model towards unseen data set . Furthermore, they did not address statistical methods that might enhance prediction, such as regularization and ensemble.

Figure 3. Coefficients of the PCE (Lloyd-Jones et al., 2014) [25]

	White			African American		
	Coefficient	Individual Example Value	Coefficient × Value†	Coefficient	Individual Example Value	Coefficient × Value†
Women (Example: 55 years of age with total cholesterol 213 mg/dL, HDL-C 50 mg/dL, untreated systolic BP 120 mm Hg, nonsmoker, and without diabetes)						
Ln Age (y)	−29.799	4.01	−119.41	17.114	4.01	68.58
Ln Age, Squared	4.884	16.06	78.44	N/A	N/A	N/A
Ln Total Cholesterol (mg/dL)	13.540	5.36	72.59	0.940	5.36	5.04
Ln Age × Ln Total Cholesterol	−3.114	21.48	−66.91	N/A	N/A	N/A
Ln HDL-C (mg/dL)	−13.578	3.91	−53.12	−18.920	3.91	−74.01
Ln Age × Ln HDL-C	3.149	15.68	49.37	4.475	15.68	70.15
Ln Treated Systolic BP (mm Hg)	2.019	—	—	29.291	—	—
Ln Age × Ln Treated Systolic BP	N/A	N/A	N/A	−6.432	—	—
Ln Untreated Systolic BP (mm Hg)	1.957	4.79	9.37	27.820	4.79	133.19
Ln Age × Ln Untreated Systolic BP	N/A	N/A	N/A	−6.087	19.19	−116.79
Current Smoker (1=Yes, 0=No)	7.574	0	0	0.691	0	0
Ln Age × Current Smoker	−1.665	0	0	N/A	N/A	N/A
Diabetes (1=Yes, 0=No)	0.661	0	0	0.874	0	0
Individual Sum			−29.67			86.16
Mean (Coefficient × Value)	N/A	N/A	−29.18	N/A	N/A	86.61
Baseline Survival	N/A	N/A	0.9665	N/A	N/A	0.9533
Estimated 10-y Risk of Hard ASCVD	N/A	N/A	2.1%	N/A	N/A	3.0%
Men (Example: 55 years of age with total cholesterol 213 mg/dL, HDL-C 50 mg/dL, untreated systolic BP 120 mm Hg, nonsmoker, and without diabetes)						
Ln Age (y)	12.344	4.01	49.47	2.469	4.01	9.89
Ln Total Cholesterol (mg/dL)	11.853	5.36	63.55	0.302	5.36	1.62
Ln Age × Ln Total Cholesterol	−2.664	21.48	−57.24	N/A	N/A	N/A
Ln HDL-C (mg/dL)	−7.990	3.91	−31.26	−0.307	3.91	−1.20
Ln Age × Ln HDL-C	1.769	15.68	27.73	N/A	N/A	N/A
Ln Treated Systolic BP (mm Hg)	1.797	—	—	1.916	—	—
Ln Untreated Systolic BP (mm Hg)	1.764	4.79	8.45	1.809	4.79	8.66
Current Smoker (1=Yes, 0=No)	7.837	0	0	0.549	0	0
Ln Age × Current Smoker	−1.795	0	0	N/A	N/A	N/A
Diabetes (1=Yes, 0=No)	0.658	0	0	0.645	0	0
Individual Sum			60.69			18.97
Mean (Coefficient × Value)	N/A	N/A	61.18	N/A	N/A	19.54
Baseline Survival	N/A	N/A	0.9144	N/A	N/A	0.8954
Estimated 10-y Risk of Hard ASCVD	N/A	N/A	5.3%	N/A	N/A	6.1%

Another commonly seen model in the US is the General CVD FRS which only incorporates sex-specific equations among the white population and predicts total CVD instead of hard ASCVD in PCE[25]. Apart from the US-based models, the European population also has its ASCVD risk prediction models such as the SCORE[26] and the PROCAM[27] model. Additionally, some countries also have their version of ASCVD risk prediction models, such as the RAMA-EGAT score from Thailand[28]. Despite the availability of these rigorously developed risk scores, the general population-based risk score among PLHIV mostly leads to misestimation due to the difference in the range of traditional ASCVD risk factors and the lack of HIV-specific predictors in the general population model[29–31].

1.3.3 ASCVD Risk Prediction Model among PLHIV

The limitation from the general population model led to the development of a PLHIV-specific ASCVD risk score. The D:A:D risk score uses the main traditional predictors from the FRS, CD4+ lymphocyte count, and cumulative use of cART for risk quantification[32,33]. Despite having more HIV-specific predictors, the D:A:D model does not predict as well in other PLHIV population[13]. Two main limitations are that the D:A:D population mainly derived from European and Australian population and the use of traditional predictor from FRS instead of from PCE[29]. More recently, the Centers for AIDS Research Network of Integrated Clinical Systems came up with data-driven risk estimation based on PCE, which turns out to be moderately calibrated but did not improve the prediction[30]. Despite the availability of adjudicated MI outcomes, the study lacks other adjudicated ASCVD outcomes such as stroke and the relatively short mean follow-up duration[30].

1.3.4 Current Recommendation for ASCVD Risk Prediction for PLHIV

The most recent scientific statement on ASCVD risk prediction among PLHIV has proposed a guideline for clinical practice based on the available literature[13]. For PLHIV in the age range of 40 – 75 years, use of PCE is justifiable with alternative use of D:A:D or FRS. In addition to the equation, HIV-related CVD risk-enhancing factors need to be considered. These factors include: a history of prolonged HIV viremia or delayed cART initiation, low current or nadir CD4 count, HIV treatment failure or low adherence, metabolic syndrome, fatty liver, and hepatitis C virus coinfection[13]. The presence of any of these risk-enhancers increases the risk upward by 1.5 – 2 folds, particularly for a history of prolonged viremia, delayed cART initiation, and low CD4 count[13]. Determination of high risk should be based on the following cut-off value: 7.5% for the PCE, 3.5% for D:A:D, and 10% for FRS[13]. Furthermore, other

12

characteristics can be incorporated into this final decision, such as a family history of ASCVD[13]. A more systematically developed PLHIV-based model is still needed to refine the model.

1.3.5 Subclinical ASCVD Risk Prediction among PLHIV

The slow progression of subclinical atherosclerosis can be a target of preventive intervention. However, there is no predictive model that is specifically built to target this upstream process. Existing studies are more likely to apply established ASCVD risk scores to predict this early process[34]. However, this approach also led to misestimation of risk for subclinical atherosclerosis among PLHIV even with the use of PLHIV derived models such as the D:A:D risk score[35,36]. This misestimation is potentially related to the difference in atherosclerosis processes among PLHIV and the change in outcome from clinical events to the subclinical development of the disease

1.4 Mortality Risk prediction tool among PLHIV – the VACS Index

The clinical spectrum of ASCVD includes death from ASCVD events such as fatal MI and fatal stroke in hard ASCVD outcomes. This broad clinical spectrum leads to overlapping predictions between ASCVD and mortality. One of the most widely used mortality risk predictions among PLHIV is the VACS index. The VACS index was developed and validated among PLHIV in 2013 (VACS index version 1)[37]. The index was developed among veteran patients and validated among other North American and European populations. The scoring was later revised in 2019 to incorporate more predictors to increase its discriminatory property (VACS index version 2)[38]. The VACS index offers a clinically related interpretation of

mortality and is available as an online accessible risk calculator[39]; therefore, it has been used in several settings to predict mortality among PLHIV.

The VACS index incorporates age, CD4 counts, HIV RNA viral load, hemoglobin level, platelets, aspartate transaminase, alanine transaminase, creatinine, and viral hepatitis C infection to make a prediction primarily for all-cause mortality[37]. The index is accurate in other populations such as those newly initiating cART, those who have been on cART for the first year, those highly treated, and among young military recruits[37,40]. The prediction is also acceptable regardless of sex, age, HCV infection status, and HIV viral suppression status[37,40,41]. Additional analysis was done that supports its predictability for cause-specific mortality, including cardiovascular mortality among PLHIV, which is comparable to its prediction for all-cause mortality[42]. Furthermore, the index is associated with a variety of biomarkers of inflammation which includes cystatin C, TNF-α, IL-6, sCD14, sCD163, and D-dimer[43–45]. Beyond the markers of inflammation, the index is related to the Chronic Immune Activation and Senescence score[46].

From a modeling perspective, the VACS Index is a 5-year mortality risk prediction model. The index was developed from the PLHIV in the VACS cohort study (4,932 PLHIV with 656 deaths) and validated in the Antiretroviral Therapy Cohort Collaboration (3146 PLHIV with 86 deaths) both from 2000 to 2007[37]. The derived cox model for 5-year risk score contained categorical age, CD4 cell counts, HIV-1 RNA viral load, hemoglobin, FIB-4, GFR, and hepatitis C infection[37]. The VACS index translates into 5-year mortality based on parametric Gamma regression model from combined NA-ACCORD and VA participants[38].

The VACS Index illustrated a clinically-oriented and simple risk score developed among PLHIV. However, the risk score depended only on the categorical version of each predictor and

no interaction terms. Additionally, regularization or ensemble strategies to improve prediction were not implemented during the model development process.

Figure 4. The interface of the VACS Index calculator (available at https://vacs.med.yale.edu/calculator/IC) [39]

1.5 Aims, Hypothesis, and Rationale

Risk quantification of CVD-related outcomes among PLHIV still needs further refinement through models that better capture the difference in risk factors and underlying mechanism among PLHIV compared to that of the general population. The range of the ASCVD clinical spectrum leads to different risk prediction depending on the natural history of disease that might offer a window of opportunity for prevention effort. In this dissertation, we will explore the prediction of subclinical atherosclerosis, non-HIV mortality, and all-cause mortality and compare the predictive performance of the newly derived model with corresponding established clinical risk prediction tools in terms of aims and hypotheses. A brief rationale for these prediction models follows.

(1.) Develop a predictive model for new focal atherosclerotic plaques among PLHIV from MWCCS incorporating additional potential predictors and compare its' predictive performance with PCE. We hypothesize that the newly derived model will have better predictive performance for new focal atherosclerotic plaques than the PCE.

Rationale:

(1.1) The PCE did not accommodate the range of values for traditional ASCVD risk factors among the PLHIV since it was derived from the general population.

(1.2) The PCE lacks risk factors specific to PLHIV.

(1.3) Despite being the standard prediction model, the PCE was originally designed to predict clinical events rather than subclinical ASCVD events.

(1.4) Modern estimation methods that might enhance predictive accuracy such as shrinkage or penalization method were not implemented in the modeling process for the PCE

(2.) Develop a predictive model for all-cause mortality among PLHIV from MWCCS and compare the predictive performance with the VACS Index and the PCE. We hypothesize that the newly derived model will have comparable predictive performance to the VACS Index and better performance than the PCE.

Rationale:

(2.1) Regarding the outcome of prediction, the PCE did not encompass all-cause mortality but only coronary death and fatal stroke.

(2.2) The VACS Index simplifies its predictors to categorical formats without interaction terms to encourage transparency; however, this simplification might not maximize performance.

(2.3) Both the PCE and the VACS Index did not include machine learning methods to improve predictive accuracy.

(3.) Develop a predictive model for non-HIV mortality among PLHIV from MWCCS and compare the predictive performance with the VACS Index and the PCE. We hypothesize that the newly derived model will have better predictive performance than the VACS Index and the PCE.

Rationale:

(1) Both the PCE and the VACS Index were not specifically designed to use in the presence of competing risk.

(2) No boosting approach that could help improve prediction was implemented in PCE or VACS Index

1.6 Reference

1   Karim QA, Vermund S, Singh U, Baxter C, Karim SSA. Acquired Immuno-Deficiency Syndrome (AIDS). In: *Oxford Textbook of Public Health*.; 2018.

2   Hutchinson AB, Farnham PG, Dean HD, Ekwueme DU, Del Rio C, Kamimoto L, *et al.* **The economic burden of HIV in the United States in the era of highly active antiretroviral therapy: Evidence of continuing racial and ethnic differences**. *J Acquir Immune Defic Syndr* 2006; **43**:451–457.

3   Ritchwood TD, Bishu KG, Egede LE. **Trends in healthcare expenditure among people living with HIV/AIDS in the United States: evidence from 10 Years of nationally representative data**. doi:10.1186/s12939-017-0683-y

4   Ward T, Sugrue D, Hayward O, Mcewan P, Anderson S-J, Lopes S, *et al.* Estimating HIV Management and Comorbidity Costs Among Aging HIV Patients in the United States: A Systematic Review. ; 2020. www.jmcp.org (accessed 9 Jun2021).

5   UNAIDS. Fact Sheet - World AIDS Day 2020. 2020.https://www.unaids.org/en/resources/fact-sheet (accessed 4 Jan2021).

6   CDC. **HIV Surveillance Report, 2018 (Updated)**. 2020; **31**.https://www.cdc.gov/hiv/library/reports/hiv-surveillance/vol-31/index.html#acknowledgments

7   CDC. Diagnose and Treat to Save Lives: Decreasing Deaths Among People with HIV. 2020.https://www.cdc.gov/hiv/statistics/deaths/index.html (accessed 4 Jan2021).

8   Kelly JA. **Ten Things We Need to Do to Achieve the Goals of the End the HIV Epidemic Plan for America**. *J Acquir Immune Defic Syndr* 2019; **82**:S94–S98.

9   Messer C, Baker J. **Ending the HIV epidemic by 2030 requires immediate action**. *JAAPA* 2019; **32**:8–9.

10   Palella FJ, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, *et al.* **Mortality in the Highly Active Antiretroviral Therapy Era**. *JAIDS J Acquir Immune Defic Syndr* 2006; **43**:27–34.

11   Feinstein MJ, Bahiru E, Achenbach C, Longenecker CT, Hsue P, So-Armah K, *et al.*

Patterns of Cardiovascular Mortality for HIV-Infected Adults in the United: 1999 to 2013. *Am J Cardiol* 2016; **117**:214–220.

12   Wong ND, Amsterdam EA, Blumenthal RS. *ASPC Manual of Preventive Cardiology*. New York: Demos Medical Publishing, LLC; 2015.

13   Feinstein MJ, Hsue PY, Benjamin LA, Bloomfield GS, Currier JS, Freiberg MS, *et al.* **Characteristics, Prevention, and Management of Cardiovascular Disease in People Living with HIV: A Scientific Statement from the American Heart Association**. *Circulation* 2019; **140**:e98–e124.

14   Shah A, Stelzle D, Lee KK, Beck E, Alam S, Clifford S, *et al.* **Global Burden of Atherosclerotic Cardiovascular Disease in People Living With HIV: Systematic Review and Meta-Analysis**. *Circulation* Published Online First: 2018.https://www.ahajournals.org/doi/pdf/10.1161/CIRCULATIONAHA.117.033369

15   Mdodo R, Frazier EL, Dube SR, Mattson CL, Sutton MY, Brooks JT, *et al.* **Cigarette Smoking Prevalence Among Adults With HIV Compared With the General Adult Population in the United States Cross-sectional Surveys**. doi:10.7326/M14-0954

16   Xu Y, Chen X, Wang K. Global prevalence of hypertension among people living with HIV: a systematic review and meta-analysis. J. Am. Soc. Hypertens. 2017; **11**:530–540.

17   Worm SW, De Wit S, Weber R, Sabin CA, Reiss P, El-Sadr W, *et al.* **Diabetes Mellitus, Preexisting Coronary Heart Disease, and the Risk of Subsequent Coronary Heart Disease Events in Patients Infected With Human Immunodeficiency Virus: The Data Collection on Adverse Events of Anti-HIV Drugs (D:A:D Study)**. Published Online First: 2009. doi:10.1161/CIRCULATIONAHA.108.790857

18   Sabin C, Worm S, Weber R, Reiss P, El-Sadr W, Dabis F, *et al.* **Use of nucleoside reverse transcriptase inhibitors and risk of myocardial infarction in HIV-infected patients enrolled in the D:A:D study: a multi-cohort collaboration**. *Lancet* 2008; **371**:1471–1426.

19   Sabin CA, Reiss P, Ryom L, Phillips AN, Weber R, Law M, *et al.* **Is there continued evidence for an association between abacavir usage and myocardial infarction risk in individuals with HIV? A cohort collaboration**. Published Online First: 2016. doi:10.1186/s12916-016-0588-4

20   Marcus JL, Neugebauer R, Neugebauer RS, Leyden WA, Chao CR, Xu L, *et al.* **Use of Abacavir and Risk of Cardiovascular Disease Among HIV-Infected Individuals**. Published Online First: 2015. doi:10.1097/QAI.0000000000000881

21   Elion RA, Althoff KN, Zhang J, Moore RD, Gange SJ, Kitahata MM, *et al.* **Recent abacavir use increases risk for Types 1 and 2 myocardial infarctions among adults with HIV HHS Public Access**. *J Acquir Immune Defic Syndr* 2018; **78**:62–72.

22   Kingsley LA, Deal J, Jacobson L, Budoff M, Witt M, Palella F, *et al.* **Incidence and progression of coronary artery calcium (CAC) in HIV-infected and HIV-uninfected men HHS Public Access**. *AIDS* 2015; **29**:2427–2434.

23   Fitch K V, Srinivasa S, Abbara S, Burdo TH, Williams KC, Eneh P, *et al.* **Noncalcified Coronary Atherosclerotic Plaque and Immune Activation in HIV-Infected Women**. Published Online First: 2013. doi:10.1093/infdis/jit508

24   Zanni M V, Abbara S, Lo J, Wai B, Hark D, Marmarelis E, *et al.* **Increased Coronary Atherosclerotic Plaque Vulnerability by Coronary Computed Tomography**

Angiography in HIV-Infected Men Program in Nutritional Metabolism.
doi:10.1097/QAD.0b013e32835eca9b

25   Lloyd-Jones DM, Braun LT, Ndumele CE, Smith SC, Sperling LS, Virani SS, *et al.* **Use of Risk Assessment Tools to Guide Decision-Making in the Primary Prevention of Atherosclerotic Cardiovascular Disease: A Special Report From the American Heart Association and American College of Cardiology**. *J Am Coll Cardiol* 2019; **73**:3153–3167.

26   Conroy R, Pyo¨ra¨la K, Fitzgerald A, Sans S, Menotti A, De Backer G, *et al.* **Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project**. *Eur Heart J* 2003; **24**:987–1003.

27   Assmann G, Cullen P, Schulte H. **Simple Scoring Scheme for Calculating the Risk of Acute Coronary Events Based on the 10-Year Follow-Up of the Prospective Cardiovascular Münster (PROCAM) Study**. *Circulation* Published Online First: 2002.http://www.circulationaha.org (accessed 28 Jan2021).

28   Vathesatogkit P, Woodward M, Tanomsup S, Ratanachaiwong W, Vanavanan S, Yamwong S, *et al.* **Cohort Profile: The electricity generating authority of Thailand study**. *Int J Epidemiol* 2012; **41**:359–365.

29   Thompson-Paul AM, Lichtenstein KA, Armon C, Palella FJ, Skarbinski J, Chmiel JS, *et al.* **Cardiovascular Disease Risk Prediction in the HIV Outpatient Study**. *Clin Infect Dis* Published Online First: 2016. doi:10.1093/cid/ciw615

30   Feinstein MJ, Nance RM, Drozd DR, Ning H, Delaney JA, Heckbert SR, *et al.* **Assessing and refining myocardial infarction risk estimation among patients with human**

immunodeficiency virus: A study by the Centers for AIDS Research Network of Integrated Clinical Systems. *JAMA Cardiol* 2017; **2**:155–162.

31   Triant VA, Perez J, Regan S, Massaro JM, Meigs JB, Grinspoon SK, *et al.* **Cardiovascular Risk Prediction Functions Underestimate Risk in HIV Infection**. *Circulation* Published Online First: 2018. doi:10.1161/CIRCULATIONAHA.117.028975

32   Friis-Møller N, Thié R, Reiss P, Weber R, D' A, Monforte A, *et al.* **Predicting the risk of cardiovascular disease in HIV-infected patients: the Data collection on Adverse Effects of Anti-HIV Drugs Study**. *Eur J Cardiovasc Prev Rehabil* 2010; **17**:491–501.

33   Friis-Møller N, Ryom L, Smith C, Weber R, Reiss P, Dabis F, *et al.* **An updated prediction model of the global risk of cardiovascular disease in HIV-positive persons: The Data-collection on Adverse Effects of Anti-HIV Drugs (D:A:D) study**. *Eur J Prev Cardiol* 2016; **23**:214–223.

34   Topel ML, Shen J, Morris AA, Mheid A, Sher S, Dunbar SB, *et al.* **Comparisons of the Framingham and Pooled Cohort Equation Risk Scores for Detecting Subclinical Vascular Disease in Blacks Versus Whites**. *Am J Cardiol* 2018; **121**:564–569.

35   Parra S, Coll B, Aragonés G, Marsillach J, Beltrán R, Rull A, *et al.* **Nonconcordance between subclinical atherosclerosis and the calculated Framingham risk score in HIV-infected patients: relationships with serum markers of oxidation and inflammation**. *HIV Med* 2010; **11**:225–231.

36   Serrano-Villar S, Estrada V, Gómez-Garre D, Ávila M, Fuentes-Ferrer M, San RJ, *et al.* **Diagnosis of subclinical atherosclerosis in HIV-infected patients: higher accuracy of the D:A:D risk equation over Framingham and SCORE algorithms**. *Eur J Prev*

Cardiol 2014; **21**:739–748.

37   Tate JP, Justice AC, Hughes MD, Bonnet F, Reiss P, Mocroft A, *et al.* **An internationally generalizable risk index for mortality after one year of antiretroviral therapy**. *AIDS* 2013; **27**:563–572.

38   Tate JP, Sterne JAC, Justice AC, Tate J. **Albumin, white blood cell count, and body mass index improve discrimination of mortality in HIV-positive individuals**. *AIDS* Published Online First: 2019. doi:10.1097/QAD.0000000000002140

39   VACS Calculator. 2021.https://vacs-apps2.med.yale.edu/calculator (accessed 28 Jan2021).

40   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**. *J Acquir Immune Defic Syndr* 2013; **62**:149–163.

41   Gustafson DR, Shi Q, Holman S, Minkoff H, Cohen MH, Plankey MW, *et al.* **Predicting death over 8 years in a prospective cohort of HIV-infected women: the Women's Interagency HIV Study**. *BMJ Open* 2017; **7**:13993.

42   **Reply to Chow et al**. *Clin Infect Dis* 2012; **55**:751–752.

43   Justice AC, Freiberg MS, Tracy R, Kuller L, Tate JP, Goetz MB, *et al.* **Does an Index Composed of Clinical Data Reflect Effects of Inflammation, Coagulation, and Monocyte Activation on Mortality Among Those Aging With HIV?** doi:10.1093/cid/cir989

44   Mooney S, Tracy R, Osler T, Grace C. **Elevated biomarkers of inflammation and coagulation in patients with HIV are associated with higher framingham and VACS risk index scores**. *PLoS One* 2015; **10**. doi:10.1371/journal.pone.0144312

45   Williams B, Livak B, Bahk M, Keating SM, Adeyemi OM. **Short Communication: SCD14 and SCD163 Levels Are Correlated with VACS Index Scores: Initial Data from the Blunted Immune Recovery in CORE Patients with HIV (BIRCH) Cohort**. *AIDS Res Hum Retroviruses* Published Online First: 2016. doi:10.1089/aid.2015.0012

46   Duffau P, Wittkop L, Lazaro E, le Marec F, Cognet C, Blanco P, *et al.* **Association of immune-activation and senescence markers with non-AIDS-defining comorbidities in HIV-suppressed patients**. *AIDS* 2015; **29**:2099–2108.

METHODS

2.1 Study Population

2.1.2 MACS

In 1983, gay and bisexual men were recruited to enroll in a longitudinal study of HIV infection in the US, contributing tremendously to the body of knowledge in HIV/AIDS[1]. The MACS, the first and largest study of its kind, has followed more than 7,300 participants twice a year, collecting data through in-depth interviews, conducting physical examinations, and storing biological specimens for subsequent testing[1,2]. The study has been carried out at four sites which are UCLA, Northwestern University in Chicago, the University of Pittsburgh, and John Hopkins University in Baltimore[1,2]. Apart from building a biological repository, the original study objective ranges from studying the natural history of HIV/AIDS, risk factors of AIDS, psychological–behavioral characteristics and interactions, temporal and spatial influence on manifestation and determinants of infectious process[1–3]. The timely adjustment to the changing epidemics allows MACS to study several key research question that PLHIV encountered; for instance, the effect of HIV, cART, and age on inflammation and immune dysfunction and non-AIDS outcomes, which are cancer, cardiovascular, liver, metabolic, neurologic, psychologic, and renal disease[1–3].

2.1.2 WIHS

Ten years after the initiation of the MACS, a paralleled study was initiated in response to the need to study HIV/AIDS in women[4–6]. Established in 1993, the Women's Interagency HIV Study (WIHS) has enrolled approximately 5,000 women across the US to address the impact of HIV/AIDS among women[4–6]. The WIHS included scientifically rigorous design, update-to-date protocol, representative samples of women in the US, and standardized specimen collection. The results in this highly successful cohort study has contribute many unique

observations of HIV/AIDS in women such as reproductive health, clinical outcomes including

non-communicable diseases and the effectiveness of cART among women[3–6]. The

comparative study timeline between MACS and WIHS is shown in the following Figure 1.

Figure 1. Study Timeline between MACS and WIHS (MWCCS, 2019) [26]

*In April 1995, 2255 HIV-negative men were administratively censored from further follow-up, per an NIH-decision
** The MACS was initiated in 1983 and the WIHS was initiated in 1994.
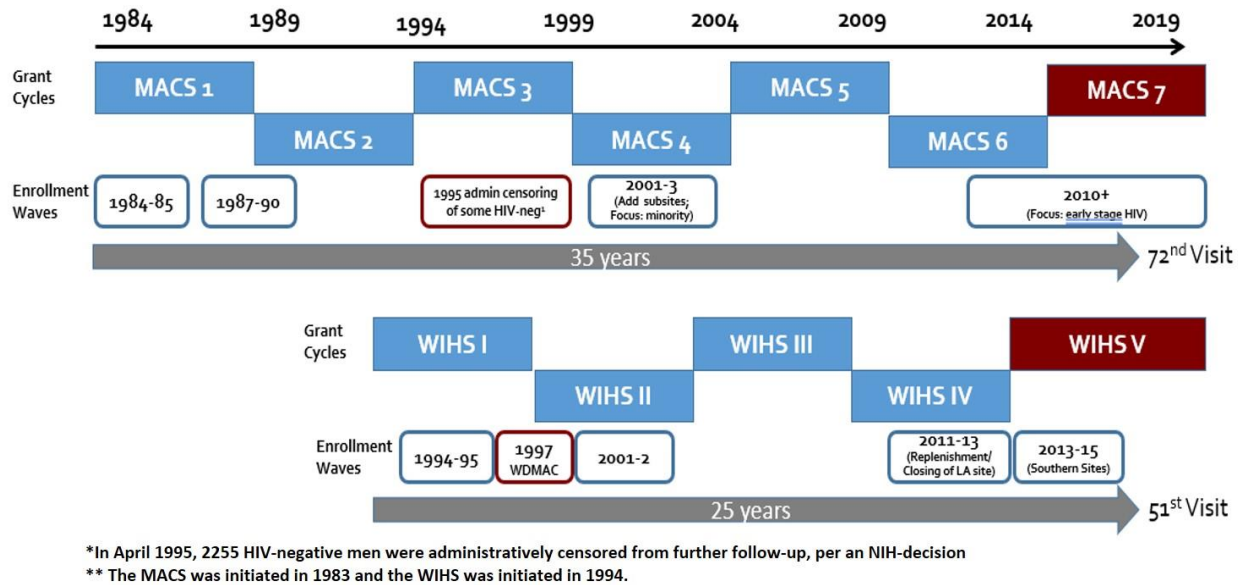
2.1.3 MWCCS

Since the introduction of cART, PLHIV have a longer life expectancy leading to the rise

of non-AIDS comorbidities and non-AIDS deaths among PLHIV[7,8]. This new focus resulted

in the change in the funding sources for both MACS and WIHS from the NIAID to the NHLBI

under the NIH[3]. This new funding scheme led to the harmonization of both cohorts into the

MACS/WIHS Combined Cohort Study (MWCCS) to continuously explore the impact of chronic

health conditions including heart, lung, blood, and sleep disorder among PLHIV[3,9].

Furthermore, this scientific endeavor also covers co-occuring conditions among PLHIV,

including mental health, neurological illness, diabetes, kidney failure, liver disease, and

cancers[3,9].

2.1.4 Vascular Sub-study among MWCCS

Among both MACS and WIHS, a vascular sub-study was initiated in 2004[10]. All WIHS participants were eligible for participation, while MACS participants were limited to those reporting no history of coronary heart disease[10]. MACS also excluded participants aged < 40 years and weight > 300 pounds since the sub-study in the MACS included coronary artery calcium measurement that required CT scan[11,12]. The sub-study was added to the proposed protocol starting from visit between 2004 and 2006 through 2013[10] with follow up of this particular sub-study occurred every $2 - 3$ years. Apart from the standardized data collection based on the semi-annual core follow-up, additional demographic information, clinical and laboratory variables were obtained from the additional sub-studies[10]. For imaging study, the sub-study from both cohorts underwent B-mode carotid artery ultrasound to evaluate 6 locations in the right carotid artery: the near and far wall of the common carotid artery, carotid bifurcation, and the internal carotid artery[13,14]. This measurement provided information on changes in IMT in millimeter and the focal plaque formation in any of the walls based on localized IMT more than 1.5 mm[10]. The CAC score from coronary CT scan was only available in the MACS study[11,12].

2.2 Predictor variables

Predictor variables studied in this dissertation include traditional ASCVD risk factors, inflammatory biomarkers, radiographic variables, risk enhancers, and predictor variables from the VACS index.

2.2.1 Traditional ASCVD Risk Factors

Some predictors repeatedly appear among these risk assessment tools despite various cardiovascular risk prediction tools. These variables are sometimes called traditional risk factors for ASCVD instead of novel biomarker and radiographic assessment. In this dissertation, these variables are referred to as the existing variables in the PCE since it is the standard risk assessment tool for the US population. These predictor variables are age, sex, total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), systolic blood pressure (SBP), smoking status, antihypertensive medication, diabetes mellitus, and race[15]. Unsurprisingly, some of these variables also appear in other ASCVD risk assessment tools such as the Framingham Risk Score, the Reynold score, and SCORE equation[15].

2.2.2 PLHIV-specific risk factors

Variables related to routine care of PLHIV have been incorporated into cardiovascular prediction as seen in the D:A:D model[16,17]. These variables include cART and HIV blood testing, including CD4+ cell counts, CD8+ cell counts, and HIV viral load. The cART variables were related to potential side effects on the cardiometabolic factors associated with each cART. For the blood test, these laboratory results are associated with the virus's activity, reflecting the amount of chronic inflammation in the body.

2.2.3 Inflammatory Biomarkers

Since PLHIV have an altered inflammatory profile due to chronic inflammation, additional predictors directly focus on inflammation are needed to capture this phenomenon better. Several inflammatory markers have been explored, such as c-reactive protein (CRP), interleukin-6 (IL-6), soluble cluster of difference-14 (sCD14), and macrophage activation markers. Despite the availability and promising predictive capability of these variables, predictive performance changes tend to be limited. For example, although the sCD14 turned out to be statistically significant in predicting carotid plaque formation, the improvement in C-statistics is only 0.01[18].

2.2.4 Novel Imaging Modalities

Novel imaging modalities can yield more insight into risk prediction. One of the most robust imaging predictors in CVD development to date is the CAC score[15]. The underlying mechanism of this powerful predictive capability is the proportional relationship between the degree of coronary calcification and the overall burden of coronary atherosclerosis[15]. The addition of CAC to the traditional risk model results in significant improvement in model performance , and the increase in predictive performances supersede other novel risk markers[15]. In the general population, CAC has been incorporated in the US multi-society guideline for an adult with uncertain risk decision in selected adults for statin initiation[19]. Despite a very well demonstrated improvement in risk prediction, the utility of CAC towards ASCVD risk prediction has not yet been fully understood, especially among particular populations such as PLHIV.

2.2.5 Other ASCVD risk factors

In the absence of biomarker and imaging that is well-studied as ASCVD predictors among PLHIV, it is justifiable to apply other potential ASCVD risk factors to the PLHIV population. Some of these risk predictors are based on ASCVD risk enhancers from the 2018 ACC/AHA cholesterol clinical guideline, such as body weight, height, waist circumference, diastolic blood pressure, blood glucose, low-density lipoprotein cholesterol (LDL-C), and triglyceride level[20]. We also included statin use in this category. For WIHS, self-reported perception of menopause is also included as a potential risk factor[20].

2.2.6 Predictors for the VACS Index

Some of the predictors overlap with previous categories for the VACS Index, including age, sex, race, body weight, height, CD4+ level, HIV-1 RNA viral load, serum creatinine, and hepatitis C infection. Apart from these predictors, the VACS Index requires hemoglobin level, AST, ALT, platelet count, and white blood cell count for risk prediction[21].

2.3 Outcome variables

Based on the wide clinical spectrum of ASCVD outcomes, the three ASCVD related outcomes in this dissertation are new focal atherosclerosis plaque, non-HIV mortality, and all-cause mortality.

2.3.1 New Focal Atherosclerosis Plaque

Subclinical atherosclerosis can be evaluated by ultrasonographic assessment of the internal carotid artery[22]. Ultrasound techniques can identify intima-media thickness (IMT) and atherosclerotic plaque[22]. While IMT is visualized by echography on both walls of the common carotid artery in a longitudinal image by a double-line pattern, plaques are focal structures situated on the inner surface of the arterial lumen with the additional thickness of at least 0.5 mm or 50% of the surrounding IMT value, or > 1.5 mm as measured from the intima-lumen interface to the media-adventitia interface as illustrated in Figure 2[22]. These two options of ultrasonographic evaluation have a different protocol that needs to be specified in clinical care and research[22]. Despite the same evaluation method, carotid IMT and plaque are different phenotypes with different levels of vascular risk[22]. Since atherosclerotic plaque has a higher risk of an atherosclerotic event than IMT, carotid plaque will be the measurement for subclinical atherosclerosis in this study. In the vascular sub-study of both MACS and WIHS, plaque measurement takes place over six locations: the near and far wall of the common carotid artery, carotid bifurcation, and internal carotid artery[10]. As opposed to IMT, atherosclerotic plaque in the sub-study was evaluated only at two-time points for each participant with the definition based on the area of localized intima-media thickness > 1.5mm[10]. Any participant without plaque on any of the six walls at baseline evaluation but later develop focal plaque on any of the walls will be classified as having new focal atherosclerotic plaque over the follow-up period[10].

Figure 2. Drawn representation of the carotid tree, with the plaque and IMT measurement according to Mannheim consensus: (1) thickness > 1.5 mm; (2) lumen encroaching > 0.5 mm; (3, 4) > 50% of the sur- rounding IMT value (Touboul PJ et al., 2012) [22]
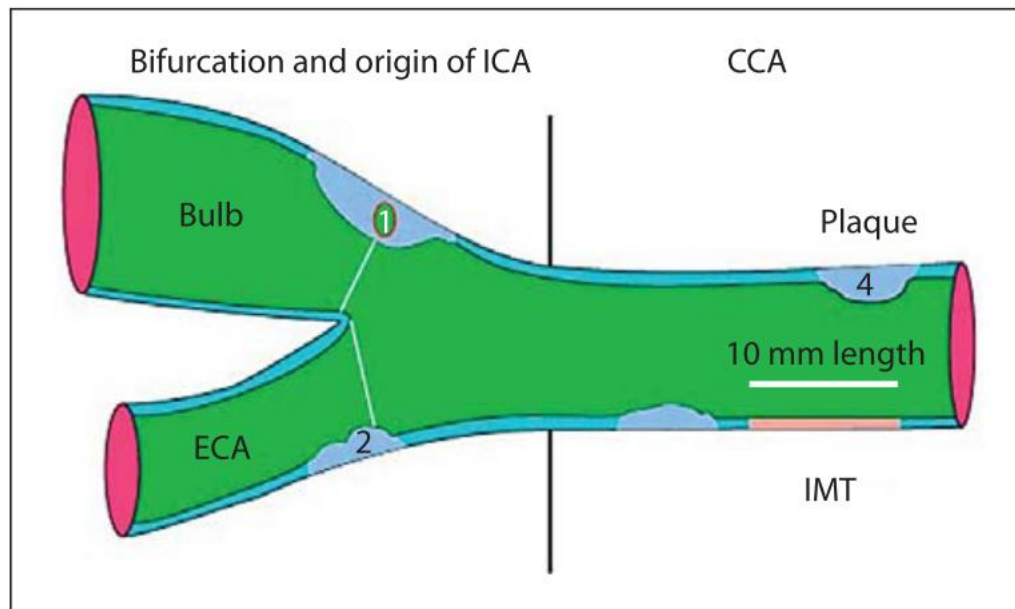
2.3.2 Non-HIV mortality and All-cause Mortality

Non-HIV mortality is based on the underlying cause of death from death certificates and ICD-10 coding. A death that did not include the ICD-10 code for HIV as the underlying cause of death (B20-B24) will be considered. However, deaths due to external causes will be excluded: accident, intentional self-harm, or assault (V01-Y89), or psychoactive substances (F11-F16, F18-F19). This exclusion focuses on deriving natural occurring death among the PLHIV as a surrogate of ASCVD-specific mortality[23].

Lastly, all-cause mortality is defined as death from any cause among the participants during both cohorts' follow-up period. This death information is ascertained both by the cohort observations and linkage with the National Death Index[23].

2.4 Statistical analysis

2.4.1 Regression models

We choose a regression model based on the outcome. In Chapter 3, a binary outcome is presented to indicate the presence or absence of a new atherosclerotic plaque at seven years. Thus, we will use a logistic regression model for this prediction. In Chapter 4, a time-to-event outcome is presented, reflecting survival from all-cause mortality. We will use the Cox proportional hazard model to predict the risk. In Chapter 5, a time-to-event outcome with competing risk is presented. The primary outcome was non-HIV death, so all HIV-related death competes for the risk of death among the participants. Competing risk analysis can be carried out by either the cause-specific hazard model or the Fine and Gray sub-distribution hazard model. The difference is that the prior is more appropriate for the etiology question. At the same time, the latter is more applicable to estimate the covariate effect on the absolute risk of outcome over time[24–26]. Consequently, we will utilize the Fine and Gray sub-distribution hazard model in Chapter 5. For the Cox model and the sub-distribution hazard model, the proportional hazard assumptions still apply and are tested by either residual plots or a statistical test of interaction between each covariate and time.

2.4.2 Non-linearity and transformation

Each regression model comes with assumptions of linearity; however, non-linearity is commonly found in the relationship between each predictor and the outcome. We tested for non-linearity by plots between the predictors and the outcomes on a linear scale. We considered transformations such as adding the quadratic and cubic term, logarithmic transformation, exponential change, or square root of the predictor. However, if none of the applicable transformations improved the linear relationship, we left the variable as is (untransformed).

34

2.4.3 Machine learning variable selection methods

The modern prediction model required more sophisticated statistical methods and frameworks. The earlier automated techniques, such as stepwise regression, suffered several limitations, including (1) the instability of variable selection, (2) the biased estimation of coefficients, (3) the misspecification of variability and exaggeration of p-values, (4) the possibility of worse predictive performance than from a model with all available predictors[27].

Since we want the most accurate result, predictive accuracy is the primary concern for model selection. Thus, a strategy for model selection needs to comply with this concern. In this dissertation, we will use algorithm that applies cross-validation for this task. The process starts with dividing the data into random subsets. We then leave one subset as a testing data set while use the rest of the subsets to train the model. We eventually repeat this process several times with different training and testing subsets.

2.4.3.1 Concept of shrinkage and penalization

We develop a prediction model from a sample at hand, aiming to apply risk prediction in other comparable populations. However, this generalization might not hold. This vital threat is the problem of overfitting. From the statistical standpoint, overfitting is a curse of dimensionality from fitting a model with too many covariates. A paralleled viewpoint from statistical modeling would be to use an overt effective degree of freedom in the modeling process. Causes of overfitting are model uncertainty and parameter uncertainty. Since models are not predefined, the information of data under study essentially drives the specification of our model, forming the basis of model uncertainty. Parameter uncertainty is a consequence of uncertainty in the effects of each predictor leading to a too extreme prediction from a model[27].

35

Figure 3. Conceptual depiction of overfitting: the left plot is overly fit to the current data and might not predict well for other value of predictor outside this data set (Yadlowsky et al., 2018)[28]
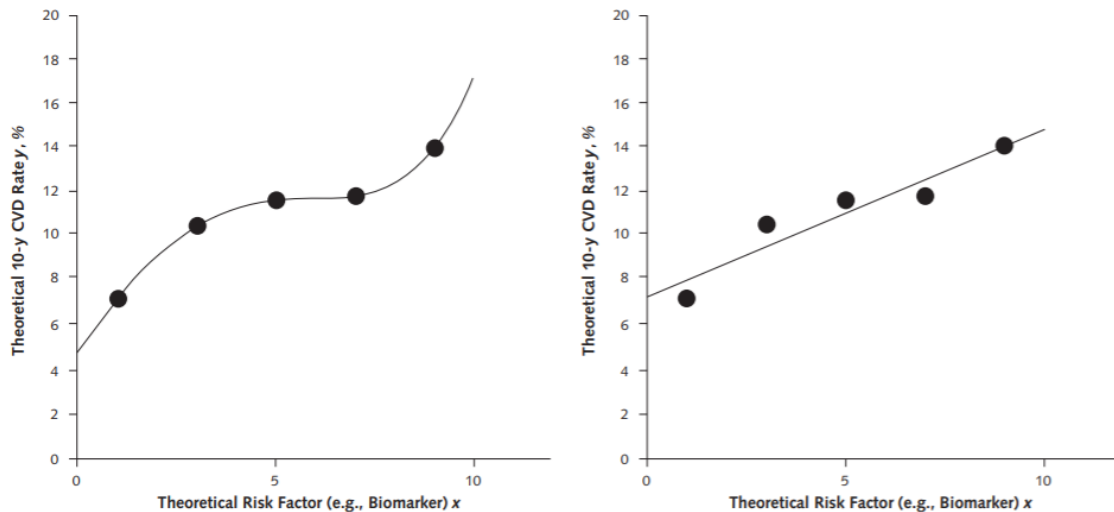
Overfitting leads to optimism and testimation bias. Optimism is the overt performance achieved in a new subject from the population that gave rise to the model. However, the performance decreases when applying the model to a new population. Testimation bias is the overestimation of the effect of predictors because of the selection of effect that withstood statistical testing. Both consequences are not favorable for predictions. The solution to this overfitting problem is shrinkage[27].

From a linear regression perspective, prediction aims to minimize the mean squared error, the square distance between the observed and predicted outcomes. In other words, we can improve prediction if we can shrink the projection towards the average[29]. Therefore, we can reduce the mean squared error for the future subject using slightly biased regression coefficients. This sacrifice of the unbiased estimate can potentially lead to a significant gain in statistical

efficiency and avoidance of overfitting. The traditional approach is to apply shrinkage to the regression coefficient after the initial fit hence the name "shrinkage after estimation[27]."

An alternative method to the traditional shrinkage is a penalized estimation. Penalization use penalty factor during the estimation of regression coefficients. This procedure penalizes the larger value of standardized coefficients in the fitting process and preferentially retains the smaller values. Thus, penalization is often viewed as "shrinkage during estimation[27]." Three related penalized regression are the ridge, the lasso, and the elastic net regression. These three-technique utilize different penalty factors: the L2 penalty term for the ridge, the L1 penalty term for the lasso, and a mix of L1 and L2 for the elastic net[29]. In this dissertation, we will utilize the elastic net regularization in chapters 3 and 4.

As shown below, elastic net regression has the penalty factors as a mix of both L1 and L2 penalty factors.

$$\hat{\beta} = argmin||y - X\beta||^2 + \lambda_2||\beta||^2 + \lambda_1||\beta||_1$$

We can see two penalty terms: the linear and the quadratic ones. The linear (L2) penalty term $(||\beta||_1)$ is from the lasso while the quadratic (L1) penalty term $(||\beta||^2)$ is from the ridge. This mixture gives the elastic net regularization a compromise between the ridge and the lasso regression. When applied separately, the ridge regression only shrinkage the coefficient with more penalty on larger coefficient while the lasso results in variable selection by shrinking coefficients to zero (sparsity selection) and tends to penalize more uniformly. Despite offering more interpretation models based on the concept of parsimony, the lasso faces limitation in setting with large number of predictors and small sample size where it only selects at most n predictors. Additionally, the lasso fails in the group selection scenario where it appears to pick

just one predictor among the correlated group and ignores the others. The addition of L1 penalty terms to the L2 in elastic net regression fixed this shortcoming, borrowing the strength of both regularization techniques[29].

Chapters 3 and 4 will utilize the elastic net regression based on the R package glmnet and glmnetUtils, which solve the following problem[30,31].

$$\min_{\beta_0,\beta} \frac{1}{N} \sum_{i=1}^{N} w_i l(y_i, \beta_0 + \beta^T x_i) + \lambda \left[(1-\alpha)\|\beta\|_2^2/2 + \alpha\|\beta\|_1 \right]$$

The elastic net penalty term is the right-hand-sided term in this specification—two tuning parameters, lambda (λ) and alpha (α), are involved. The λ covers the whole grid of penalty; therefore, it controls the overall strength of the penalty. The α bridges the gap between the lasso ($\alpha = 1$) and the ridge ($\alpha = 0$). We will err on the side of parsimony and choose the largest lambda value such that the error is within one standard error of minimum (the 1 SE rule)[32].

2.4.3.2 Concept of ensemble methods – boosting

Ensemble methods improve predictive performance by using information from various models. These methods create a population of models either by (1) training the same algorithm to different versions of a data set such as in bagging and boosting or (2) training qualitatively different models on the same data set such as in Bayesian model averaging or Super learner[33].

In this dissertation, chapter 5 utilizes boosting during model development. The boosting trains models on a subset of data sequentially. Each iteration results in predictive errors that will improve the classifier in the next iteration. The resulting model is an average across each model

from the algorithmic process. While the main focus of boosting is to avoid overfitting, each boosting algorithm also incorporates additional features that avoid model preselection[33].

The R package Coxboost help carried out the likelihood-based boosting for the Fine-Gray model in Chapter 5[34]. This process is the offset-based boosting approach for estimating Cox proportional hazard model. The package can specify the cause of interest to yield the Fine-Gray model for competing risk with the boosting method. Each boosting step incorporates the next boosting step as penalized partial likelihood estimation for an update by one covariate in each boosting step. The updating results in many estimated coefficients of zero; hence it is called a lasso-like approach for this comparable sparsity selection. The tuning parameter in this algorithm is the number of boosting steps which can be determined by cross-validation[34].

Figure 4. Conceptual illustration of boosting (Modified from Bi et al., 2019)[33]

The final output classifier is a weighted average from each iteration, with higher weights given to the classifiers with higher predictive accuracy.
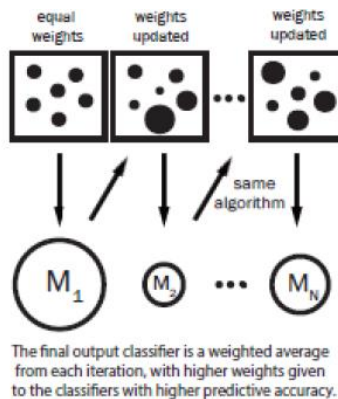
2.4.4 Model performance evaluation

We will evaluate the models in chapters 3 to 5 based on the following property: discrimination, calibration, and net reclassification.

2.4.4.1 Discrimination

The measure of discrimination addresses the extent to which a model predicts a higher probability of having an event among who will and who will not have the event[35]. This evaluation is achievable by computation of c-statistics or the graphical display of ROC curve[35]. Either of them has the same basis of incorporating sensitivity and specificity of a test into a single measure. For a binary outcome, a standard cut-off value of 0.7 for both value of c-statistics and AUC of the ROC curve is often considered. One common finding from related studies is that this measure is hardly increased by incorporating any additional predictors if traditional predictors in the model are strongly predictive of the outcome.

2.4.4.2 Calibration

Calibration is the extent to which a model correctly estimates the risk[35]. This measure usually is regarded as the most critical aspect of a predictive model. A poorly calibrated model will usually underestimate or overestimate the outcome of interest. Calibration can be evaluated either by graphical methods such as plotting the decile of predicted and observed event rate or by statistical methods including chi-square test and expected-to-observed probability ratio[35]. In some cases, recalibration might further help improve prediction from the model; however, this is not guaranteed, and it also depends on the availability of information needed for this process.

2.4.4.3 Net Re-classification

Risk stratification of each individual might also be re-classified by updating the risk prediction model compared to the existing one. NRI answers how much more frequently appropriate reclassification occurs than inappropriate reclassification using the new model[35]. Thus, this measurement helps illustrate the improvement in risk classification. One limitation of NRI in this setting is the need for a specific cut-point for classification, which tends to be arbitrary. Thus, a variety of cut points and sensitivity analysis is warranted. We can further classify NRI to NRI of an event (NRI_e) and NRI of non-event (NRI_{ne}). The NRI_e is the net proportion of events assigned a higher risk or risk category. The NRI_{ne} is the net proportion of non-event given a lower risk or risk category[36].

2.5 Sample size estimation

Notably, sample size estimation is difficult in the algorithmic-based approach[37]. Many recommendations on sample size estimation for machine learning are available, and most are discipline-specific[37]. This dissertation will utilize the rule of thumbs for model development that requires at least ten samples per predictor as a minimum rule of thumb[27]. Since we will apply both quantitative and qualitative evaluation in the testing data sets, we will focus on the possible number of events in the training data sets for this estimation.

For Chapter 3, a prior study from MWCCS reported approximately 164 new focal plaques for new focal plaque consisting of 94 in MACS and 70 in WIHS[10]. The training data with 70:30 splitting will contain around 115 events for the MWCCS, 65 events for the MACS, and 50 for the WIHS. Therefore, about 11, 6, and 5 predictors should be allowable to be in the

MWCCS, MACS, and WIHS models. Notably, this prior study only uses reported history of the cardiovascular event as exclusion criteria[10]. Thus, fewer events might be encountered with more restricted exclusion criteria.

For Chapter 4, we approximated all-cause mortality after 2000 from the difference between the total number of deaths in the CCS dossier file and mortality data before 2002 among PLHIV[38,39]. For MACS, there were 2148 total deaths and 1745 deaths before 2002, both among PLHIV. This difference results in 403 all-cause deaths among the MACS. For the WIHS, there were 1240 total deaths and 590 deaths before 2002, both among PLHIV. This difference leads to 650 all-cause deaths among the WIHS. This approximation led to a total of 1053 deaths among MWCCS. From the 70% data retained in the training portion, we should have at most 70 predictors in the MWCCS model, 28 predictors in the MACS model, and 45 predictors in the WIHS models.

For Chapter 5, there were 2148 total deaths, 1745 deaths before 2002, and non-HIV to all-cause death was $37/72 \approx 0.513$ among MACS. These figures result in approximately 207 non-HIV deaths in the MACS. For the WIHS, there were 1240 total deaths, 590 deaths before 2002, and non-HIV to all-cause death was $90/219 \approx 0.411$. These numbers result in about 267 non-HIV deaths in the WIHS. This approximation led to a total of 474 non-HIV death in MWCCS. From the 70% of total data in the training portion, we should have at most 33 predictors in the MWCCS model, 14 predictors in the MACS model, and 18 predictors in the WIHS models.

Despite the estimations above, it is worth noting that it is still possible to obtain fewer participants from the database due to additional exclusion criteria than the reference articles, the cross-validation process, the time period selected for model development, and variation in reported number used in the estimation.

2.6 Reference

1   Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR. **The Multicenter AIDS Cohort Study: Rationale, Organization, and Selected Characteristics of the Participant**. *Am J Epidemiol* 1987; **126**:310–318.

2   Detels R, Jacobson L, Margolick J, Martinez-Maza O, Muñoz A, Phair J, *et al.* **The multicenter AIDS Cohort Study, 1983 to ...** *Public Health* 2012; **126**:196–198.

3   MACS/WIHS Combined Cohort Study. https://statepi.jhsph.edu/mwccs/about-mwccs/ (accessed 27 Jan2021).

4   Adimora AA, Ramirez C, Benning L, Greenblatt RM, Kempf M-C, Tien PC, *et al.* **Cohort Profile: The Women's Interagency HIV Study (WIHS)**. *Int J Epidemiol* 2018; **47**:393-394i.

5   Bacon MC, Von Wyl V, Alden C, Sharp G, Robison E, Hessol N, *et al.* **The Women's Interagency HIV Study: an Observational Cohort Brings Clinical Sciences to the Bench**. *Clin DIiagnostic Lab Immunol* 2005; **12**:1013–1019.

6   Barkan S, Melnick S, Preston-Martin S, Weber K, Kalish L, Miotti P, *et al.* **The Women's Interagency HIV Study**. *Epidemiology* 1997.

7   Feinstein MJ, Bahiru E, Achenbach C, Longenecker CT, Hsue P, So-Armah K, *et al.* **Patterns of Cardiovascular Mortality for HIV-Infected Adults in the United in 2013; however, CVD mortality increased from 307 to HHS Public Access**. *Am J Cardiol* 2016; **117**:214–220.

8   Palella FJ, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, *et al.* **Mortality in**

the Highly Active Antiretroviral Therapy Era. *JAIDS J Acquir Immune Defic Syndr* 2006; **43**:27–34.

9   Souza G D', Golub ET, Gange SJ. **The Changing Science of HIV Epidemiology in the United States**. *Am J Epidemiol* 2019; **188**. doi:10.1093/aje/kwz211

10   Hanna DB, Post WS, Deal JA, Hodis HN, Jacobson LP, Mack WJ, *et al.* **HIV Infection Is Associated With Progression of Subclinical Carotid Atherosclerosis**. *Clin Infect Dis* 2015; **61**:640–650.

11   Kingsley LA, Cuervo-Rojas J, Muñoz A, Palella FJ, Post W, Witt MD, *et al.* **Subclinical coronary atherosclerosis, HIV infection and antiretroviral therapy: Multicenter AIDS Cohort Study**. *AIDS* Published Online First: 2008. doi:10.1097/QAD.0b013e328306a6c5

12   Post WS, Budoff M, Kingsley L, Palella FJ, Witt MD, Li X, *et al.* **Associations between HIV Infection and Subclinical Coronary Atherosclerosis: The Multicenter AIDS Cohort Study (MACS)**. doi:10.7326/M13-1754

13   Kaplan RC, Kingsley LA, Gange SJ, Benning L, Jacobson LP, Lazar J, *et al.* **Low CD4+ T-cell count as a major atherosclerosis risk factor in HIV-infected women and men**. *AIDS* 2008; **22**:1615–1624.

14   Hodis HN, Mack WJ, Lobo RA, Shoupe D, Sevanian A, Mahrer PR, *et al.* Estrogen in the Prevention of Atherosclerosis. ; 2001. https://annals.org (accessed 27 Jan2021).

15   Wong ND, Amsterdam EA, Blumenthal RS. *ASPC Manual of Preventive Cardiology*. New York: Demos Medical Publishing, LLC; 2015.

16   Friis-Møller N, Thiébaut R, Reiss P, Weber R, Monforte AD, De Wit S, *et al.* **Predicting the risk of cardiovascular disease in HIV-infected patients: the data collection on adverse effects of anti-HIV drugs study.** *Eur J Cardiovasc Prev Rehabil* 2010; **17**:491–501.

17   Friis-Møller N, Ryom L, Smith C, Weber R, Reiss P, Dabis F, *et al.* **Original scientific paper An updated prediction model of the global risk of cardiovascular disease in HIV-positive persons: The Data-collection on Adverse Effects of Anti-HIV Drugs (D:A:D) study**. *Eur J Prev Cardiol* 2016; **23**:214–223.

18   Hanna DB, Lin J, Post WS, Hodis HN, Xue X, Anastos K, *et al.* **Association of Macrophage Inflammation Biomarkers With Progression of Subclinical Carotid Artery Atherosclerosis in HIV-Infected Women and Men**. *J Infect Dis J Infect Dis* 2017; **215**:1352–61.

19   Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, *et al.* **2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines**. *J Am Coll Cardiol* 2019; **73**:e285–e350.

20   Feinstein MJ, Hsue PY, Benjamin LA, Bloomfield GS, Currier JS, Freiberg MS, *et al.* **Characteristics, Prevention, and Management of Cardiovascular Disease in People Living with HIV: A Scientific Statement from the American Heart Association**. *Circulation* 2019; **140**:e98–e124.

21   VACS Index Information and VACS Calculator.

https://medicine.yale.edu/intmed/vacs/vacsresources/vacsindexinfo/ (accessed 27 Jan2021).

22   Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, *et al.* Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011). Cerebrovasc. Dis. 2012; **34**:290–296.

23   Hanna DB, Moon J-Y, Haberlen SA, French AL, Palella FJ, Gange SJ, *et al.* **Carotid Artery Atherosclerosis is associated with Mortality in HIV-Positive Women and Men**. *AIDS* 2018; **32**:2393–2403.

24   Austin PC, Lee DS, Fine JP. **Introduction to the Analysis of Survival Data in the Presence of Competing Risks**. *Circulation* 2016; **133**:601–609.

25   Austin PC, Fine JP. **Practical recommendations for reporting Fine-Gray model analyses for competing risk data**. Published Online First: 2017. doi:10.1002/sim.7501

26   Fine JP, Gray RJ. **A Proportional Hazards Model for the Subdistribution of a Competing Risk**. *J Am Stat Assoc* 1999; **94**:496–509.

27   Steyerberg EW. Statistics for Biology and Health Clinical Prediction Models A Practical Approach to Development, Validation, and Updating Second Edition. ; 2019. http://www.springer.com/series/2848 (accessed 14 Apr2020).

28   Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min YI, Basu S. **Clinical implications of revised pooled Cohort equations for estimating atherosclerotic cardiovascular disease risk**. *Ann Intern Med* 2018; **169**:20–29.

29   James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning*. ; 2013.

30   Friedman J, Hastie T, Tibshirani R, Narasimhan B, Tay K, Simon N, *et al.* Package
      "glmnet." ; 2021. doi:10.18637/jss.v033.i01

31   Package "glmnetUtils" . ; 2021. https://github.com/hongooi73/glmnetUtils (accessed 17
      May2021).

32   Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining,
      Inference, and Prediction.* ; 2017. doi:10.1111/j.1751-5823.2009.00095_18.x

33   Bi Q, Goodman KE, Kaminsky J, Lessler J. **What is Machine Learning? A Primer for
      the Epidemiologist**. *Am J Epidemiol* Published Online First: 2019.
      doi:10.1093/aje/kwz189

34   Binder H. **Package ' CoxBoost .'** 2015.

35   Lloyd-Jones DM. **Risk Prediction in Cardiovascular Medicine Cardiovascular Risk
      Prediction Basic Concepts, Current Status, and Future Directions**. *Circulation*
      Published Online First: 2010. doi:10.1161/CIRCULATIONAHA.109.849166

36   Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. **Net Reclassification
      Indices for Evaluating Risk-Prediction Instruments: A critical review**. *Epidemiology*
      2014; **25**:114–121.

37   Wiemken TL, Kelley RR. Machine learning in epidemiology and health outcomes
      research. Annu. Rev. Public Health. 2019; **41**:21–36.

38   MACS/WIHS CCS Dossier. ; 2019. https://mwccs.org/policies__publications/ (accessed 6
      Apr2020).

39   Hessol NA, Kalinowski A, Benning L, Mullen J, Young M, Palella F, *et al.* **Mortality**

among Participants in the Multicenter AIDS Cohort Study and the Women's Interagency HIV Study. *Clin Infect Dis* 2007; **44**:287–294.

Manuscript #1

Prediction model for a new subclinical atherosclerotic plaque from cardiovascular risk factors among people living with HIV

Abstract

Background: People living with HIV (PLHIV) experience more cardiovascular diseases due to increased life expectancy from the combined antiretroviral therapy. The subclinical phase of the disease is an opportunity for prevention; however, a risk prediction model for subclinical atherosclerosis among PLHIV has not been established.

Methods: The included participants were from the vascular sub-study of the MWCCS without a history of coronary heart disease, including angina, myocardial infarction, coronary revascularization, stroke, or use of a cardiac pacemaker. We fitted the elastic net logistic regression separately by sex and for the whole MWCCS using new atherosclerotic plaque over seven years as the outcome. We benchmarked the new model against the pooled cohort equation (PCE) for discrimination, calibration, and net reclassification.

Results: A total of 754 participants, 263 males, and 491 females, were included in this analysis. The prevalence of new atherosclerotic plaque was around 11%. The new models appear to have higher discrimination than the PCE; however, their 95%CI overlapped. Expected-to-observed ratio indicated that the new model had better mean calibration while the PCE tended to underestimate the risk. Calibration plots suggested that only the new models from the MWCCS had comparable calibration with the PCE. The net reclassification supported that the newly derived model reclassifies participants to the higher risk groups.

Conclusion: Our model building approach led to model with better mean calibration and sensible risk reclassification as compared to the PCE. Further refinement is needed to improved discrimination and pattern on the calibration plots of the new models.

3.1 Background

The introduction of combined antiretroviral therapy (cART) in 1996 has significantly changed the natural history of HIV infection[1]. This breakthrough resulted in a significant decrease in AID-related deaths, a downward trend in opportunistic infections, and lengthened longevity among people living with HIV (PLHIV)[1]. This increased survival allows PLHIV to experience non-communicable diseases (NCDs), including atherosclerotic cardiovascular disease (ASCVD), even earlier than the general population counterpart. The accelerated rate of ASCVD is concerning due to the sudden nature of its clinical manifestation[1].

The long-standing subclinical stage of ASCVD provides an opportunity for preventive strategies. In routine clinical practice, ASCVD risk assessments have been the cornerstone of preventive cardiology[2]. Therefore, detecting the early stage of ASCVD based on the risk calculator might allow us to use this window of opportunity to halt or even reverse the disease process. Despite the available risk assessment tools, applying these risk calculators to predict subclinical atherosclerosis among PLHIV appears to misestimate the risk[3]. This misestimation potentially stems from the general population-derived risk score among PLHIV, which has a different range of predictor values and possibly different mechanisms of atherosclerosis formation. Even the D:A:D score, a PLHIV-based score, suffered this miscalculation[4]. Since the scoring was initially designed to predict clinical events, applying those calculators towards a subclinical event likely needs some adjustments.

Furthermore, these established prediction models use unpenalized regression models[5]. Implementation of shrinkage or penalization technique might offer a more accurate prediction[6,7]. These modern estimation techniques require a different approach to model

building, such as the cross-validation process to obtain models with desired level of performance measures[6,7].

To address this gap in the literature, we proposed using the Multicenter AIDS Cohort Study (MACS)/ WIHS (Women's Interagency HIV Study) Combined Cohort Study (MWCCS) data to derive a PLHIV-specific risk score for predicting subclinical atherosclerosis formation by penalized regression and compared this newly derived risk score with the Pooled Cohort Equations (PCE) regarding discrimination, calibration, and net reclassification.

3.2 Methods

3.2.1 Study Population

Participants were from the MWCCS. Details on study design are described elsewhere[8,9]. In short, participants in each study had semi-annual follow-up visits and underwent parallel detailed examinations and structured interviews. Vascular sub-study in MACS/WIHS began in 2004 with a history of coronary heart disease as the main exclusion criteria[10]. MACS also excluded participants less than 40 years and weight more than 300 pounds to enable coronary artery calcium measurement by computerized tomography (CT) scan[11]. Additional exclusion criteria for this current study were self-reported history of coronary heart disease including angina, myocardial infarction, coronary revascularization; stroke, or use of a cardiac pacemaker at the enrollment of the vascular sub-study.

The vascular sub-study had baseline visits between 2004 and 2006 and follow-up visits every two to three years until their last follow-up visit around 2011 and 2013[10]. Additional demographic, clinical, and laboratory were collected during the semi-annual visit[10].

Participants also underwent high-resolution B-mode carotid artery ultrasound to measure both intima media thickness and focal atherosclerotic plaque on any of the six walls of the right carotid artery[12]. The scan for plaque detection only occurred at the baseline visit and the last follow-up visit for all participants[10]. For MACS participants, twice cardiac scans were carried out by electron beam tomography or multidetector CT[11]. In this analysis, we included only those with complete follow-up assessment for the subclinical atherosclerosis plaque ultrasound scan. This result in 263 MACS participants, 491 WIHS participants, and 754 MWCCS participants as showed in table 1. MACS appeared to have higher mean age and more proportion of white participants than WIHS. While WIHS had more percentage of diabetes, MACS had more proportion of hypertension and dyslipidemia. Most of the variables had missing values less than 5%. Only dyslipidemia (DLP), waist circumference (WC), low-density lipoprotein cholesterol (LDL), and triglyceride (TG) had more than 5% of missing values. The research protocol was approved by the Institutional Review Board in University of California, Los Angeles (IRB# 20-001446) and the MWCCS Executive Committee (Project# X20051).

3.2.2 Predictors

Twenty-four potential predictor variables were included in this analysis. The choice of the variables was based on the traditional cardiovascular practice and PLHIV-specific literature on predictors for ASCVD[1,2,13]. We imputed missing data by median (continuous variables) or mode (categorical variables). These predictors were classified into the four domains, as listed below. Each predictor was presented with their study abbreviation and assigned numerical codes of each level (categorical variables) or unit of measurement (continuous variables).

3.2.2.1 Traditional ASCVD risk factors

We included the following traditional ASCVD risk factors: : age (AGE, years), sex (SEX; male = 1, female =0), race (RACE; African American = 1, White and others = 0), smoking status (SMKGRP; current smoker = 1, non or past smoker = 0), diabetes mellitus (DM; yes = 1, no = 0), hypertension (HTN; yes = 1, no = 0), antihypertensive treatment (HTNRX; received treatment = 1, not received treatment = 0), systolic blood pressure (SBP; mmHg), total cholesterol (TC; mg/dL), and high-density lipoprotein cholesterol (HDL; mg/dL).

3.2.2.2 Other ASCVD risk factors

Apart from the traditional ASCVD predictors, other potential ASCVD predictors were included in this domain: dyslipidemia (DYSLIP; yes = 1, no = 0), statin medication use (STATIN; yes = 1, no = 0), body mass index (BMI; kg/m^2), waist circumference (WC; cm), and low-density lipoprotein cholesterol (LDL; mg/dL) and triglyceride level (TG; mg/dL). For WIHS participants, we later excluded BMI due to high proportion of missing values. Moreover, we initially considered self-report menopausal status among WIHS participants; however, this variable was later removed due to a high percentage of missing values.

3.2.2.3 PLHIV-specific risk factors

We considered HIV-related blood tests and cART exposure in our model. HIV-related blood tests were CD4+ cell count (CD4; cell/mm^3), CD8+ cell count (CD8; cell/mm^3), and HIV RNA viral load (VL; copies/ ml). The cART exposure was cumulative years of protease inhibitor (PI; years), non-nucleotide/nucleoside reverse transcriptase inhibitor (NNRTI; years), and Abacavir (ABC; years).

3.2.2.4 Inflammatory biomarkers

In addition to PLHIV-specific risk factors, other blood tests might indicate inflammatory response but not be specific to HIV pathophysiology; therefore, we initially considered the following two blood tests as inflammatory biomarkers: c-reactive protein (CRP; mg/dL) and interleukin-6 (IL-6; pg/dL) in this group. However, we later excluded IL-6 due to a high proportion of missing values. Among WIHS participants, CRP results were obtained as weighted mean from individually available CRP testing.

3.2.2.5 Novel imaging modalities

We obtained the Coronary Artery Calcification (CAC) Score from the geometric mean of the Agatston scores[11]. The CAC was only available among the MACS participants.

3.2.3 Outcomes

A new atherosclerotic lesion was the outcome of the study. The vascular sub-study evaluated this disease progression approximately over seven years by ultrasound assessment on the six walls of the carotid artery[12].

3.2.4 Statistical Analysis: model development and evaluation

We split the initial data into training (2/3 of the data) and testing (1/3 of the data) data sets. In the training data sets, we derived separate models for MACS, WIHS, and MWCCS. This strategy led to three new models: MACS sex-specific model, WIHS sex-specific model, and MWCCS model. We would later compare these new models to the benchmarking model in the testing data sets. All the statistical analyses were done in R version 4.1.1[14].

Each potential predictor underwent the following process in the model building stage. We coded categorical variables into a separate level such that SEX0 = 1 is female while SEX1 = 1 is male. Then, interaction terms between age, sex, race with all other covariates were included. We investigated the potential non-linearity of each continuous predictor by the shape of each predictor and logits of the outcome adjusted for age, sex, and race. For continuous variable, cubic, quadratic, square root, logarithmic, and exponential transformations were considered. Additionally, we explored the specific transformation of log(CAC+1) for the CAC score among MACS participants.

The primary regression model was the logistic regression model. We implemented the elastic net regularization to fine-tune the regression coefficients (β). The method has the following penalty factor as specified in the glmnet package[15].

$$\lambda[(1 - \alpha)||\beta||_2^2/2 + \alpha||\beta||_1]$$

This penalty involves two tuning parameters: lambda (λ) and alpha (α). While the lambda controlled the overall strength of this penalty, the alpha regulated the trade-off between two penalty terms. The two penalties terms are the L1 quadratic penalty term from the ridge regularization and the L2 linear penalty term from the lasso regularization. Thus, the elastic net embraces these two related techniques resulting in shrinkage and variable selection[6]. We explored the lambda and alpha range based on cross-validation targeting at the highest area under the curve (AUC) for prediction from the glmnetUtils package [16], an extension of the glmnet package. We selected the final models based on the number of retained non-zero coefficients and the estimated AUC from the cross-validation process.

We assessed model performance based on discrimination, calibration, and net reclassification in the testing data sets. The discrimination measures were receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC) using plotROC package [17]. The calibration evaluation included an expected-to-observed ratio (E:O ratio) and the calibration plot from plotCalibration function in PredictABEL package [18]. We explored the net reclassification of events (NRI_e) and the net reclassification of non-events (NRI_{ne}) against the benchmarking models based on the 7.5% and 20% cut-off points by the nribins function in nricens package[19]. Confidence intervals were generated from bootstrapping.

3.2.5 Benchmarking Model: PCE

The PCE is a sex-race-specific survival prediction model that computes the 10-yr risk of ASCVD. This risk calculator is a standard prediction model for ASCVD prevention among the general population in the US[20]. We used the PCE as a benchmarking model. Any predictor's value more extreme than the acceptable range of PCE was truncated to the limit of that predictor from the PCE.

Table 1 Baseline characteristics in of MACS, WIHS and MWCCS participants with carotid ultrasound scans both at baseline and at follow-up visits in training and testing data sets

Characteristics	MACS (n = 263)		WIHS (n = 491)		MWCCS (n = 754)	
	Training data set (n = 173)	Testing data set (n = 90)	Training data set (n = 324)	Testing data set (n = 167)	Training data set (n = 497)	Testing data set (n = 257)
Traditional ASCVD risk factors						
Age (year)	48.31 ± 5.49	48.05 ± 5.79	41.15 ± 7.98	40.36 ± 7.87	43.64 ± 7.97	43.05 ± 8.08
Race:						
Black	50 (28.90 %)	36 (40.00 %)	174 (50.88 %)	90 (53.89 %)	224 (45.07 %)	126 (49.03 %)
White	123 (71.10 %)	54 (60.00 %)	145 (44.75 %)	73 (43.71 %)	268 (53.92 %)	127 (49.42 %)
Smoking status:						
Current smoker	57 (32.95 %)	32 (35.56 %)	138 (42.59 %)	69 (41.32 %)	195 (39.24 %)	101 (39.30 %)
Non/ past smoker	115 (66.47 %)	56 (62.22 %)	186 (57.41 %)	98 (58.68 %)	301 (60.56 %)	154 (59.92%)
Diabetes mellitus:						
Yes	36 (20.81 %)	27 (30.00 %)	29 (8.95 %)	10 (6.00 %)	65 (13.08 %)	37 (14.40 %)
No	137 (79.20 %)	63 (70.00 %)	295 (91.05 %)	157 (94.00 %)	432 (86.92 %)	220 (85.60 %)
Hypertension:						
Yes	33 (19.08 %)	24 (26.67 %)	51 (15.74 %)	27 (16.17 %)	84 (16.90 %)	51 (19.84 %)
No	140 (80.92 %)	66 (73.33 %)	272 (83.95 %)	140 (83.83 %)	412 (82.90 %)	206 (80.16 %)
Antihypertensive treatment:						
Yes	31 (17.92 %)	20 (22.22 %)	58 (17.90 %)	34 (20.36 %)	89 (17.91 %)	54 (21.01 %)
No	141 (81.50 %)	70 (77.78 %)	266 (82.10 %)	133 (79.64 %)	407 (81.89 %)	203 (78.99 %)
Systolic Blood Pressure (mmHg)	124.51 ± 13.67	127.86 ± 15.72	116.96 ± 16.98	116.00 ± 16.00	119.60 ± 16.29	121.12 ± 16.64
Total Cholesterol (mg/dL)	189.04 ± 44.67	192.20 ± 50.86	176.25 ± 42.35	180.25 ± 39.08	180.70 ± 43.56	184.43 ± 43.85
HDL (mg/dL)	44.34 ± 13.17	45.50 ± 15.71	47.73 ± 17.66	47.49 ± 16.75	46.55 ± 16.30	46.79 ± 16.39
Missing: n (%)	0 (0.00 %)	0 (0.00 %)	1 (0.31 %)	0 (0.00 %)	1 (0.20 %)	0 (0.00 %)
Other ASCVD risk factors						
Dyslipidemia:						
Yes	132 (76.30 %)	74 (82.22 %)	120 (37.04 %)	58 (34.73 %)	252 (50.70 %)	132 (51.36 %)
No	29 (16.76 %)	13 (14.44 %)	200 (61.73 %)	109 (65.27 %)	229 (46.08 %)	122 (47.47 %)
Missing: n (%)	12 (6.94 %)	3 (3.33 %)	4 (1.23 %)	0 (0.00 %)	16 (3.22 %)	3 (1.17 %)

Table 1 Baseline characteristics in of MACS, WIHS and MWCCS participants with carotid ultrasound scans both at baseline and at follow-up visits in training and testing data sets (continued)

Characteristics	MACS (n = 263)		WIHS (n = 491)		MWCCS (n = 754)	
	Training data set (n = 173)	Testing data set (n = 90)	Training data set (n = 324)	Testing data set (n = 167)	Training data set (n = 497)	Testing data set (n = 257)
Statin Medication:						
Yes	52 (30.06 %)	27 (30.00 %)	35 (10.80 %)	14 (8.38 %)	87 (17.51 %)	41 (15.95 %)
No	121 (69.94 %)	63 (70.00 %)	289 (89.20 %)	153 (91.62 %)	410 (82.49 %)	216 (84.05 %)
BMI (kg/m^2)	25.64 $\pm$ 4.14	25.84 $\pm$ 3.69	NA	NA	NA	NA
Waist circumference (cm)	92.71 $\pm$ 11.43	92.81 $\pm$ 11.08	90.36 $\pm$ 15.43	93.32 $\pm$ 14.37	91.38 $\pm$ 14.15	93.13 $\pm$ 13.24
Missing: n (%)	12 (6.94 %)	6 (6.67 %)	36 (11.11 %)	19 (11.38 %)	48 (9.66 %)	25 (9.73 %)
LDL (mg/dL)	105.93 $\pm$ 36.89	113.32 $\pm$ 39.14	102.61 $\pm$ 35.69	104.84 $\pm$ 32.02	103.65 $\pm$ 36.06	107.61 $\pm$ 34.66
Missing: n (%)	26 (15.03 %)	9 (10.00 %)	1 (0.31 %)	0 (0.00 %)	27 (5.43 %)	9 (3.50 %)
Triglyceride (mg/dL)	183.00 (188.50)	143.00 (100.00)	111.50 (83.00)	114.00 (73.00)	121.00 (112.00)	119.50 (88.25)
Missing: n (%)	26 (15.03 %)	9 (10.00 %)	0 (0.00 %)	0 (0.00 %)	26 (5.23 %)	9 (3.50 %)
PLHIV-specific risk factors						
CD4 (cell/mm^3)	549.00 (322.00)	491.00 (382.75)	452.50 (339.25)	474.00 (407.00)	492.00 (349.00)	475.00 (399.00)
CD8 (cell/mm^3)	889.00 (559.00)	942.00 (540.00)	767.50 (506.50)	763.00 (422.50)	796.00 (515.50)	804.00 (486.00)
Viral Load (count/mm^3)	50.00 (509.00)	50.00 (857.00)	180.00 (4,920.00)	80.00 (8,170)	80.00 (3,220)	80.00 (5,020.00)
Duration on PI (year)	2.90 (6.00)	2.95 (7.08)	1.50 (5.25)	0.75 (4.75)	1.96 (5.50)	1.25 (5.80)
Duration on NNRTI (year)	1.49 (4.24)	1.59 (4.34)	0.25 (2.25)	0.75 (2.75)	0.75 (2.75)	1.25 (3.25)
Duration on Abacavir (year)	0.00 (2.14)	0.00 (0.00)	0.00 (1.38)	0.00 (1.25)	0.00 (1.75)	0.00 (0.59)
Inflammatory biomarkers						
CRP (mg/dL)	1.40 (2.80)	1.10 (2.83)	2.00 (4.38)	2.70 (4.28)	1.70 (3.65)	2.20 (3.98)
Novel imaging modalities						
CAC score	0.00 (6.49)	0.00 (22.34)	NA	NA	NA	NA

Figures are presented as mean $\pm$ standard deviation or median (interquartile range) for continuous variables, depending on their distribution, and count (percentage) for categorical variables.
Missing data only indicates for those with more than 5% missing values including dyslipidemia, waist circumference, LDL, and triglyceride.

Table 2 Newly derived models from the training data sets

MACS sex-specific model		WIHS sex-specific model		MWCCS model	
Predictors	Coefficients	Predictors	Coefficients	Coefficients	Coefficients
Intercept	-1.9716	Intercept	-2.1448	Intercept	-2.0124
PI	0.0269	AGE	6.7829×10^{-6}	STATIN0	-2.9192×10^{-1}
		DM0	-7.3093×10^{-4}	CD4	-6.5554×10^{-4}
		DM1	7.3014×10^{-4}	SEX0*DM1	3.8970×10^{-1}
		STATIN0	-1.3478×10^{-3}	SEX0*STATIN1	4.0312×10^{-1}
		STATIN1	1.3470×10^{-3}	SEX1*PI	5.1100×10^{-2}
		Sqrt(ABC)	2.0110×10^{-4}	SEX0*ABC	1.1495×10^{-2}
		AGE*SBP	4.4365×10^{-9}	AGE*LDL	7.1829×10^{-5}
		AGE*DM1	2.4438×10^{-5}	RACE0*SMKGRP1	1.4566×10^{-1}
		AGE*STATIN1	2.8814×10^{-5}	RACE0*DM1	3.2686×10^{-1}
		AGE*Sqrt(ABC)	7.6221×10^{-6}		
		AGE*TC	1.5561×10^{-7}		
		AGE*LDL	1.8876×10^{-7}		
		RACE0*SMKGRP1	9.4843×10^{-5}		
		RACE0*DM1	3.4128×10^{-3}		
		RACE0*STATIN1	2.9156×10^{-4}		

Abbreviation for transformations: square root (Sqrt), natural logarithm (Ln).

Table 3 Number of new subclinical atherosclerosis and total participants in testing data sets

Testing data sets	Number of new subclinical atherosclerosis	Total number of participants
MACS	17	90
WIHS	12	167
MWCCS	29	257

Table 4 Model performance of each newly derived model in testing data sets

Testing data sets	Sex-specific models				MWCCS models				PCE	
	AUC	E:O ratio	NRI_e	NRI_{ne}	AUC	E:O ratio	NRI_e	NRI_{ne}	AUC	E:O ratio
MACS	0.63	0.71	0.47	-0.50	0.53	0.71	0.47	-0.56	0.50	0.43
	(0.51,0.75)	(0.50,1.33)	(0.20,0.79)	(-0.65,-0.38)	(0.37,0.69)	(0.49,1.33)	(0.19,0.75)	(-0.70,-0.42)	(0.36,0.65)	(0.29,0.79)
WIHS	0.60	1.46	0.92	-0.90	0.61	1.42	0.83	-0.72	0.52	0.34
	(0.46,0.73)	(0.92,2.92)	(0.75,1.00)	(-0.94,-0.85)	(0.46,0.75)	(0.89, 2.89)	(0.60,1.00)	(-0.79,-0.64)	(0.38,0.67)	(0.20,0.73)
MWCCS	0.68	1.02	0.66	-0.77	0.60	1.00	0.62	-0.67	0.60	0.39
	(0.58,0.78)	(0.76,1.47)	(0.44,0.83)	(-0.82,-0.71)	(0.49,0.71)	(0.75, 1.48)	(0.40,0.81)	(-0.73,-0.58)	(0.50,0.70)	(0.28.0.59)

Abbreviation: area under receiver operating characteristic curve (AUC), expected-to-observed ratio (E:O ratio), net-reclassification of event (NRI_e), net-reclassification of non-event (NRI_{ne})

Figure 1 ROC curves in testing data sets
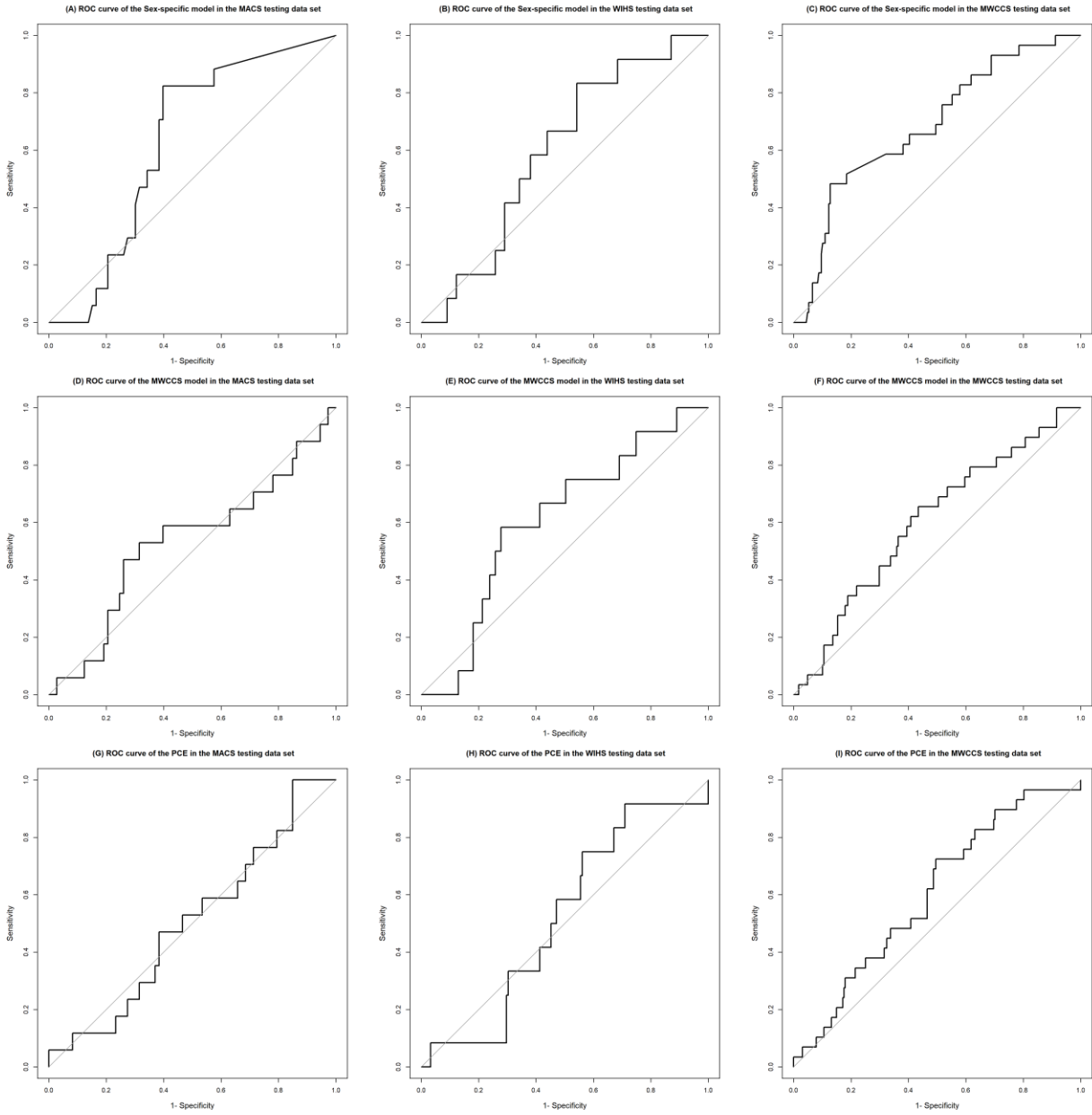
Figure 2 Calibration plots in testing data sets

(A) Calibration plot of the Sex-specific model in the MACS testing data set
(B) Calibration plot of the Sex-specific model in the WIHS testing data set
(C) Calibration plot of the Sex-specific model in the MWCCS testing data set
(D) Calibration plot of the MWCCS model in the MACS testing data set
(E) Calibration plot of the MWCCS model in the WIHS testing data set
(F) Calibration plot of the MWCCS model in the MWCCS testing data set
(G) Calibration plot of the PCE in the MACS testing data set
(H) Calibration plot of the PCE in the WIHS testing data set
(I) Calibration plot of the PCE in the MWCCS testing data set

3.3 Result

Table 2 illustrates the new models: the MACS sex-specific model, the WIHS sex-specific model, and the MWCCS model. The WIHS sex-specific model had the highest number of predictors as compared to the MACS sex-specific model or the MWCCS model. Notably, most of the coefficients were small. Furthermore, these model incorporated PLHIV-specific predictors such as CD4 or cART exposure. The sex-specific models were different between MACS and WIHS. While the MACS sex-specific model contained only one predictor, the WIHS counterpart contained many more covariates with several interactions with age and race. Additionally, we provided tuning parameters from the elastic net regularization in the Appendix (Appendix A supplementary table 1).

Table 3 presents the number of new subclinical atherosclerosis and the total number of participants in each testing data set. These figures could be translated to incidence proportion of subclinical atherosclerosis over the follow-up period of around 18.89%, 7.19%, and 11.28% among the MACS, WIHS, and MWCCS participants, respectively.

Table 4 depicts model performance measurement across testing data sets. In the MACS testing data sets, the MACS sex-specific model had higher AUC than the MWCCS model and the PCE. On the other hand, the AUC from the WIHS sex-specific model and the MWCSS model were higher than that from the PCE in the WIHS testing data set. In the MWCCS testing data set, the sex-specific models generated higher AUC than the MWCCS model and the PCE. Nevertheless, 95%CI from the new models overlapped with the PCE in each testing data set. Figure 1 provides the accompanying ROC curves of these AUC results.

Mean calibration was presented by E:O ratio in table 4. In the MACS testing data set, the MACS sex-specific model and the MWCCS model had E:O ratio of around 0.7 while this measure was 0.5 for the PCE. In the WIHS testing data set, the WIHS sex-specific model and the MWCCS model appeared to have the ratio of approximately 1.4; however, the PCE had the same figure of around 0.5 in this testing data set. In the MWCCS testing data set, the sex-specific models had the ratio of 0.68 while the MWCCS model and the PCE had a slightly lower figure of 0.6. Importantly, all the 95% CI of the sex-specific models and the MWCCS model cover 1, but all the 95%CI of the PCE situated below 1.

Calibration plots in figure 2 illustrate the agreement between predicted and observed risk in each testing data set. We could see that the sex-specific model had calibration plots which are almost vertical lines (figure 2 plots A to C). This finding means that the model had a very narrowed range of predicted risk regardless of observed risk. On the contrary, the calibration plots of the MWCCS model and the PCE model had a wider range of predicted risk as reflected from their more scattered data points (figure 2 plots D to I). These latter models tend to have better calibration as compared to the sex-specific model in each testing data set.

The net reclassification index was also included in table 4. We could see that all NRI_e from the sex-specific models and the MWCCS model was positive while the NRI_{ne} from these models were negative in each testing data set. This result reflected that the risk from the new models had positive net proportion of events assigned to the higher risk category and negative net proportion of non-events assigned to the lower risk category.

3.4 Discussion

We utilized the MWCCS data to build and evaluate the prediction models for newly formed subclinical atherosclerosis among PLHIV with the elastic net logistic regression. Although the new models had higher AUC in most of the testing data sets, their 95%CI and those from the PCE were overlapping. The E:O ratio suggested that the new models offered better mean calibration than the PCE. Specifically, this measure points towards the underestimation of risk from the PCE. Furthermore, the calibration plots illustrated that the sex-specific models had a narrower range of agreement between observed and predicted risk than the other models. The net reclassification indices supported the finding from the E:O ratio by indicating that the new models tended to assign more participants to the higher risk group as compared to the PCE.

The newly derived model exhibited some unique patterns. First, most of the coefficients were small. We expected these small magnitudes from the penalization that resulted in the shrinkage of the retained coefficients during the estimation steps. This phenomenon led to a very narrow standard deviation in risk, especially from the sex-specific models. Second, age, sex and race appeared to be important in this prediction. This observation was from the interaction terms with age and race from the WIHS sex-specific model and the interaction terms with age, sex, and race in the MWCCS model. These findings reflected the possible risk stratification with sex and race compared to the sex-race-specific cardiovascular risk model in PCE[5]. Additionally, all newly derived models included cART exposure. The inclusion of cART for cardiovascular prediction also appeared in the D:A:D model[21,22]. Nevertheless, our models incorporate statin medication that might improve this prediction from the medication perspective more than cART alone.

Despite our effort to include more PLHIV-relevant covariates, the performance of our model is not yet satisfactory. Both MACS and WIHS sex-specific models exhibited misestimation from the calibration plot. Previously, we expected a model derived from PLHIV with various covariates to overcome this struggle faced by the PCE[3] and the D:A:D[4]. The underperformance was potentially related to the very limited number of predictors in the MACS sex-specific model and the very small magnitude of coefficients in WIHS sex-specific. Only the MWCCS model had a more reasonable spread of predicted risk. Despite the misestimation in our new model, the E:O ratio and the net reclassification index still supported our belief that the PCE underestimates the risk among PLHIV.

Even though we included CAC score in our list of predictors among the MACS participants, the variable selection process from the elastic net regularization shrank its coefficients to zero. This variable selection might also have biological basis from the different mechanism of atherosclerosis among PLHIV. Although calcification is an important component of atherosclerosis formation in general population, it has been illustrated that PLHIV had higher proportion of non-calcified plaque[23,24]. Our finding from the MACS sex-specific model might be explainable by this observation as well. Therefore, further investigation by other imaging modalities such as the coronary computed tomography angiography is needed to refine the atherosclerotic plaque formation risk among PLHIV[13,25].

Our study has some limitations. First, we have limited new focal atherosclerotic plaques due to the exclusion criteria and data stratification. Second, we could not evaluate some potential predictors, such as IL-6, due to the high proportion of missing values. Third, the NRI measures relative performance that depends on arbitrary cut-off values. The result might change with other decisions on the values. Despite these limitations, our study also had many strengths. First,

67

MWCCS is a longitudinal cohort study which is an appropriate study design for prediction model development. Additionally, the cohorts also had several available predictors for model development. Second, we implemented elastic net regularization to minimize overfitting. Third, we evaluated our model in many aspects, both numerically and graphically. Thus, our work contributes additional findings to the existing body of research in prediction model development among PLHIV.

3.5 Public Health Implication

Our approach in model building led to new models with better mean calibration and appropriate risk reclassification comparing to the PCE. However, we still need more improvement in model discrimination and moderate calibration since the new models still have overlapped 95%CI of the AUC and similar pattern on calibration plots with the PCE.

3.6 Conclusion

Our efforts to develop a prediction model for new subclinical atherosclerosis among PLHIV led to improvement in the mean calibration and proper risk reclassification as compared to the PCE; nevertheless, we still haven't significantly increased discrimination based on the AUC and calibration based on the calibration plot from the new models.

3.7 Reference

1   Feinstein MJ, Hsue PY, Benjamin LA, Bloomfield GS, Currier JS, Freiberg MS, *et al.*
**Characteristics, Prevention, and Management of Cardiovascular Disease in People
Living with HIV: A Scientific Statement from the American Heart Association**.
Circulation 2019; **140**:e98–e124.

2   Wong ND, Ansterdam EA, Toth PP. *ASPC Manual of Preventive Cardiology*. Second Edi.
; 2021.

3   Parra S, Coll B, Aragonés G, Marsillach J, Beltrán R, Rull A, *et al.* **Nonconcordance
between subclinical atherosclerosis and the calculated Framingham risk score in
HIV-infected patients: relationships with serum markers of oxidation and
inflammation**. *HIV Med* 2010; **11**:225–231.

4   Serrano-Villar S, Estrada V, Gó mez-Garre D, Fuentes-Ferrer M, San Román J, Soriano
V, *et al.* **Diagnosis of subclinical atherosclerosis in HIV-infected patients: higher
accuracy of the D:A:D risk equation over Framingham and SCORE algorithms.** *Eur
J Prev Cardiol* 2014; **21**:739–748.

5   Lloyd-Jones DM, Bennett G, Coady S, D RB, Gibbons R, Greenland P, *et al.* **2013
ACC/AHA Guideline on the Assessment of Cardiovascular Risk**. *S49 (Circulation*
2014; **129**:49–73.

6   Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining,
Inference, and Prediction*. ; 2017. doi:10.1111/j.1751-5823.2009.00095_18.x

7   Friedman J, Hastie T, Tibshirani R. **Regularization paths for generalized linear models**

via coordinate descent. *J Stat Softw* 2010; **33**:1–22.

8   Detels R, Jacobson L, Margolick J, Martinez-Maza O, Muñoz A, Phair J, *et al.* **The**
       multicenter AIDS Cohort Study, 1983 to ... *Public Health* 2012; **126**:196–198.

9   Adimora AA, Ramirez C, Benning L, Greenblatt RM, Kempf M-C, Tien PC, *et al.*
       Cohort Profile: The Women's Interagency HIV Study (WIHS). *Int J Epidemiol* 2018;
       47:393-394i.

10   Hanna DB, Post WS, Deal JA, Hodis HN, Jacobson LP, Mack WJ, *et al.* **HIV Infection Is**
       Associated With Progression of Subclinical Carotid Atherosclerosis. *Clin Infect Dis*
       2015; **61**:640–650.

11   Kingsley LA, Cuervo-Rojas J, Muñoz A, Palella FJ, Post W, Witt MD, *et al.* **Subclinical**
       coronary atherosclerosis, HIV infection and antiretroviral therapy: Multicenter
       AIDS Cohort Study. *AIDS* Published Online First: 2008.
       doi:10.1097/QAD.0b013e328306a6c5

12   Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, *et al.*
       Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011).
       Cerebrovasc. Dis. 2012; **34**:290–296.

13   Patel AA, Budoff MJ. **Coronary Artery Disease in Patients with HIV Infection: An**
       Update. *Am J Cardiovasc Drugs* 2021; **21**:411–417.

14   R Core Team (2021). R: A language and environment for statistical computing. R
       Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

15   Jerome Friedman, Trevor Hastie, Robert Tibshirani (2010). **Regularization Paths for**

Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. URL https://www.jstatsoft.org/v33/i01/.

16   Hong Ooi (2021). glmnetUtils: Utilities for 'Glmnet'. R package version 1.1.8. https://CRAN.R-project.org/package=glmnetUtils

17   Michael C. Sachs (2017). **plotROC: A Tool for Plotting ROC Curves**. Journal of Statistical Software, Code Snippets, 79(2), 1-19. doi:10.18637/jss.v079.c02

18   Suman Kundu, Yurii S. Aulchenko and A. Cecile J.W. Janssens (2020). PredictABEL: Assessment of Risk Prediction Models. R package version 1.2-4. https://CRAN.R-project.org/package=PredictABEL

19   Eisuke Inoue (2018). nricens: NRI for Risk Prediction Models with Time to Event and Binary Response Data. R package version 1.6. https://CRAN.R-project.org/package=nricens

20   Lloyd-Jones DM, Bennett G, Coady S, D RB, Gibbons R, Greenland P, *et al.* **2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk**. *S49 (Circulation* 2014; **129**:49–73.

21   Friis-Møller N, Thié R, Reiss P, Weber R, D' A, Monforte A, *et al.* **Predicting the risk of cardiovascular disease in HIV-infected patients: the Data collection on Adverse Effects of Anti-HIV Drugs Study**. *Eur J Cardiovasc Prev Rehabil* 2010; **17**:491–501.

22   Friis-Møller N, Ryom L, Smith C, Weber R, Reiss P, Dabis F, *et al.* **An updated prediction model of the global risk of cardiovascular disease in HIV-positive persons: The Data-collection on Adverse Effects of Anti-HIV Drugs (D:A:D) study**. *Eur J Prev*

Cardiol 2016; **23**:214–223.

23   D'ascenzo F, Cerrato E, Calcagno A, Grossomarra W, Ballocca F, Omed P, *et al.* **High**

prevalence at computed coronary tomography of non-calcified plaques in

asymptomatic HIV patients treated with HAART: A meta-analysis. *Atherosclerosis*

2015; **240**:197–204.

24   Tarr PE, Ledergerber B, Calmy A, Doco-Lecompte T, Marzel A, Weber R, *et al.*

Subclinical coronary artery disease in Swiss HIV-positive and HIV-negative persons.

Eur Heart J 2018; **39**:2147–2154.

25   Nakanishi R, Post WS, Osawa K, Jayawardena E, Kim M, Sheidaee N, *et al.* **Multicenter**

AIDS Cohort Study (MACS) Quantitative Coronary Plaque Progression Study:

Rationale and Design. *Coron Artery Dis* 2018; **29**:23–29.

CHAPTER 4

Manuscript #2

Prediction model for all-cause mortality from cardiovascular risk factors

among people living with HIV

Abstract

Background: With a higher proportion of cardiovascular death among people living with HIV (PLHIV), a prediction model based on cardiovascular risk factors should be predictive of all-cause mortality among PHLIV.

Methods: We fitted the elastic net cox regression for the whole Multicenter AIDS Cohort Study/Women's Interagency HIV/AIDS Study Combined Cohort Study (MWCCS) in 2005 and separately by sex to predict all-cause mortality over five and ten years. We benchmarked the new model against the pooled cohort equation (PCE) and the Veterans Aging Cohort Study (VACS) Index for discrimination, calibration, and net reclassification.

Results: A total of 2,118 participants were included consisting of 1,114 male and 1,004 female participants. The risk of all-cause mortality rate was around 9 deaths per 1,000 PLHIV-year and 11 deaths per 1,000 PLHIV-year for 5-year duration and 10-year duration, respectively. Over the 5-year period, the new models had the same level of discrimination and better calibration compared to the VACS Index. Over the 10-year period, both the sex-specific models (AUC 0.68 95%CI 0.57,0.76) and MWCCS model (AUC 0.70 95%CI 0.64,0.76) had higher discrimination as compared to the PCE (AUC 0.49, 95%CI 0.42,0.55) in the MWCCS testing data set. These models also exhibit good calibration from the calibration plots and appropriate risk reclassification from the net reclassification index.

Conclusion: The new models appeared to have better predictive performance than the benchmarking models over the 10-year period. We recommend an external validation study to explore external validity of these models.

4.1 Background

The use of combined anti-retroviral therapy (cART) led to a significant decrease in deaths among people living with HIV (PLHIV)[1–3]. In the US, the substantial decrease in mortality rates among PLHIV was first noted in 1997[1]. A closer look at the recent time trend from 2010 to 2018 illustrated the continuing downward projection of all-cause mortality, with the non-HIV mortality rates greater than the HIV mortality rates[4].

Among non-HIV causes of death, cardiovascular disease (CVD) is one of the leading causes of death among PLHIV[5]. An analysis of non-HIV attributable causes of death from the US Center of Disease Control (CDC) national HIV surveillance system from 2007 to 2011 indicated that heart disease had an age-adjusted death rate of 2 per 1,000 PLHIV-year which was higher than other non-HIV causes[6]. An analysis of proportionate mortality from the CDC's Wide-Ranging Online Data for Epidemiologic Research (WONDER) from 1999 to 2013 illustrated the continuous increase in CVD proportionate mortality among PLHIV while the same relative proportion has been decreasing among the general population[5]. The same study investigated subtype of CVD, specifically ASCVD (ischemic heart disease and cerebrovascular heart disease). In this study, these two categories account for the major proportion of mortality of PLHIV, even after stratifying for sex and race.

Mortality risk prediction is integral to clinical care for both making clinical decisions and facilitating provider-patient communication. With the increase in non-HIV causes of death and the overwhelming proportion of ASCVD among CVD causes among PLHIV, it is important to investigate whether we can combine traditional risk factors with other available clinical predictor variables that are in PLHIV care to predict all-cause mortality in PLHIV.

The emerging application of machine-learning techniques is an important opportunity for prediction research[7]. One of the most important problems in prediction modeling is overfitting, that is, where the model fits too closely in one data set and therefore performs poorly elsewhere[8]. However, some statistical methods can minimize this problem. One potential tool is the use of penalization, which utilizes the machine-learning framework to generate penalty factors from cross-validation processes. These penalty factors can then be used to fine-tune newly derived models during the regression coefficient estimation steps[9]. Incorporating this modern approach to model building could potentially lead to a better performance for prediction model among PLHIV.

Therefore, we derived an all-cause mortality risk prediction model for PLHIV based on traditional ASCVD risk factors and additional clinical predictors with the use of the machine-learning framework. We compared discrimination, calibration, and net reclassification of the new models to two currently-used models: the Veteran Aging Cohort Study (VACS) Index[10–12] and the pooled cohort equations (PCE)[13].

4.2 Methods

4.2.1 Source of Data and study population

Participants were from the Multicenter AIDS Cohort Study (MACS) and Women's Interagency HIV Study (WIHS) Combined Cohort Study (MWCCS). Details on study design were described elsewhere[14–18]. In short, MACS and WIHS are multicenter longitudinal cohorts for the study of HIV among gay men and high-risk women in the US, respectively. Both

cohorts had comparable biannual visits, including interviews, physical examination, blood test, and other clinical workups.

In this analysis, we included all participants from both cohorts in 2005 as a baseline. We selected this year because the change in mortality trend occurred during the early 2000s and also because of the availability of variables in our study to calculate risk from the benchmarking models. This selection resulted in a total of 2,118 participants consisting of 1,114 from MACS and 1,004 from WIHS. Table 1 shows the baseline characteristics of our three cohorts, as randomly divided into training and testing data sets. Most predictors had their proportion of missing values less than 10% except for waist circumference (WC), low-density lipoprotein cholesterol (LDL), triglyceride (TG), and C-reactive protein (CRP). The research protocol was approved by the Institutional Review Board in University of California, Los Angeles (IRB# 20-001446) and the MWCCS Executive Committee (Project# X20051).

4.2.2 Predictors

We selected 24 potential predictor variables from the MWCCS in this analysis. The choice of the variables was based on the literature on predictors for ASCVD both in traditional practice and PLHIV-specific situation. Covariate values were obtained from the nearest visit from 2004 to 2006. Missing data were explored and imputed by median (continuous variables) or mode (categorical variables). These predictors were categorized into the four domains, as listed below. We present each predictor with their study abbreviation and assigned numerical codes of each level (categorical variables) or unit of measurement (continuous variables).

4.2.2.1 Traditional ASCVD risk factors

Traditional ASCVD risk factors comprised the following variables: age (AGE, years), sex (SEX; male = 1, female =0), race (RACE; African American = 1, White and others = 0), smoking status (SMKGRP; current smoker = 1, non or past smoker = 0), diabetes mellitus (DM; yes = 1, no = 0), hypertension (HTN; yes = 1, no = 0), antihypertensive treatment (HTNRX; received treatment = 1, not received treatment = 0), systolic blood pressure (SBP; mmHg), total cholesterol (TC; mg/dL), and high-density lipoprotein cholesterol (HDL; mg/dL).

4.2.2.2 Other ASCVD risk factors

Other potential predictors for ASCVD apart from the traditional ASCVD predictors were included in this domain: dyslipidemia (DYSLIP; yes = 1, no = 0), statin medication use (STATIN; yes = 1, no = 0), body mass index (BMI; kg/m2), waist circumference (WC; cm), low-density lipoprotein cholesterol (LDL; mg/dL) and triglyceride level (TG; mg/dL). For the WIHS, we also add self-perceived menopausal status to this list (MENO; yes = 1, no = 0).

4.2.2.3 PLHIV-specific risk factors

We incorporated blood tests and cART exposure in our model. HIV-related blood tests were CD4+ cell count (CD4; cell/mm^3), CD8+ cell count (CD8; cell/mm^3), and HIV RNA viral load (VL; copies/ ml). The cART exposure was cumulative years of protease inhibitor (PI; years), non-nucleotide/nucleoside reverse transcriptase inhibitor (NNRTI; years), and Abacavir (ABC; years).

4.2.2.4 Inflammatory biomarkers

In addition to PLHIV-specific risk factors that are part of the pathophysiologic mechanism, other blood tests might reflect inflammatory response but not be specific to HIV pathophysiology; therefore, we initially considered the following two blood tests as inflammatory biomarkers: c-reactive protein (CRP; mg/dL) and interleukin-6 (IL-6; pg/dL) in this group. However, we later excluded IL-6 due to a high proportion of missing values. Among WIHS participants, CRP results were weighted mean from individually available CRP testing.

4.2.3 Outcome

The outcome was deaths among MACS and WIHS participants from 2006 to 2015. For both cohorts, deaths were ascertained from the National Death Index and death certificates. Person time was enumerated in person-years. Loss-to-follow-up was assumed to occur between the last-know visit and the subsequent missed visit. Any participants who was lots to follow-up during the study period and later returned after 2015 received full person time during this study period.

4.2.4 Statistical analysis: model development and evaluation

The original data were randomly split into training (2/3 of the data) and testing (1/3 of the data) data sets. In the training data set, we built separate models for two different prediction periods. For 5-year projection, we use follow up data from 2006 to 2010 and for the 10-year projection we use follow up data from 2006 to 2015. Within each time period, we derived cohort specific models for MACS, WIHS, and MWCCS. This approach resulted in six newly derived models: 5-year MACS sex-specific model, 5-year WIHS sex-specific model, 5-year MWCCS

model, 10-year MACS sex-specific model, 10-year WIHS sex-specific model and 10-year

MWCCS model. These newly derived models and the two benchmarking models would later be

evaluated in the testing data set. All the statistical analyses were done in R version 4.1.1[19].

During the model-building stage, each variable in the training data set went through the

following process. We coded each categorical variable separately for each level such that SEX0

= 1 was for females and SEX1=1 was for males. Then, two-way interaction terms between age,

sex, and race with all other variables were generated. We next explored the potential non-

linearity of each continuous predictor using a generalized additive model adjusted for age, sex,

and race. Lastly, transformation of each continuous variable based on cubic, quadratic, square

root, logarithmic, and exponential terms was considered.

The primary regression model was the Cox proportional hazard model. We fine-tuned the

regression coefficients (β) by elastic net regularization. As described in glmnet package, the

penalty term is as follows [20].

$$\lambda[(1 - \alpha)||\beta||_2^2/2 + \alpha||\beta||_1]$$

Two tuning parameters, lambda (λ) and alpha (α), are used in this penalization[20]. The λ

parameter covers the whole grid; thus, it controls the degree of penalization[20]. The α parameter

acts as a weighting between two penalty terms: L1, the quadratic penalty term from the ridge

regularization, and L2, the linear penalty term from the lasso regularization[20]. Thus, the elastic

net gains advantages from both, resulting in shrinkage of coefficients and variable selection[20].

To finetune both parameters, the glmnetUtils package[21], an extension of the glmnet

package, was used in this step. We selected the final models based on the 1 SE rule [22], the

number of retained non-zero coefficients and the estimated C-index from the 10-fold cross-

validation process in the prepared training data set. The Cox proportional hazard assumption was assessed by plots between scaled Schoenfeld residuals and time.

We evaluated model performance using discrimination, calibration, and net reclassification in the testing data set. The discrimination measure was the area under the receiver operating characteristic curve (AUC) using survivalROC package [23]. The evaluation of calibration was based on calibration plot by calPlot function in pec package [24]. We compared both the net reclassification of events (NRI_e) and the net reclassification of non-events (NRI_{ne}) of each new model with the benchmarking models using the 7.5% and 20% cut-off points via nricens package [25]. These two cut-off points were taken from the ASCVD risk cut-off points from the PCE. Bootstrapping was used to generate 95% confidence interval (CI).

4.2.5 Benchmarking models: VACS Index and PCE

We included two benchmarking models for this analysis: the VACS Index and the PCE. The VACS Index is a mortality risk prediction tool based on the HIV population from US veteran databases, which generated 5-year all-cause mortality risk among PLHIV[10,11]. We calculated the VACS Index from the published model (VACS Index-I)[10,11] and obtained the 5-year mortality risk from the corresponding online risk calculator[12]. The PCE is a set of sex-race specific equations derived from pooled cohorts of the general US population[13]. The PCE yields a 10-year risk of ASCVD including coronary death, non-fatal coronary heart disease, fatal stroke, and non-fatal stroke. We calculated the predicted 10-year risk based on the publicly available details on the PCE from the 2013 American College of Cardiology/ American Heart Association (ACC/AHA) report[13]. During this process, any individual predictor with its value out of reference range of each risk models was truncated to the upper or lower limits of that model.

Table 1 Baseline characteristics in 2005 of MACS, WIHS and MWCCS participants in training and testing data sets

Characteristics	MACS (n = 1,114)		WIHS (n = 1,004)		MWCCS (n = 2,118)	
	Training data set (n = 739)	Testing data set (n = 375)	Training data set (n = 658)	Testing data set (n = 346)	Training data set (n = 1,397)	Testing data set (n = 721)
Traditional ASCVD risk factors						
Age (year)	46.41 ± 8.93	46.60 ± 9.24	45.78 ± 8.13	46.06 ± 7.27	46.12 ± 8.56	46.34 ± 8.36
Race:						
Black	238 (32.21 %)	108 (28.80 %)	321 (48.78 %)	168 (48.55 %)	559 (40.01 %)	276 (38.28 %)
White	501 (67.79 %)	267 (71.20 %)	337 (51.22 %)	178 (51.45 %)	838 (59.99 %)	445 (61.72 %)
Smoking status:						
Current smoker	273 (36.94 %)	117 (31.20 %)	291 (44.22 %)	147 (42.49 %)	564 (40.37 %)	264 (36.62 %)
Non/ past smoker	451 (61.03 %)	249 (66.40 %)	360 (54.71 %)	194 (56.07 %)	811 (58.05 %)	443 (61.44 %)
Diabetes mellitus:						
Yes	162 (21.92 %)	84 (22.40 %)	85 (12.92 %)	43 (12.43 %)	247 (17.68 %)	127 (17.61 %)
No	577 (78.08 %)	291 (77.60 %)	573 (87.08 %)	303 (87.57 %)	1,150 (82.32 %)	594 (82.39 %)
Hypertension:						
Yes	184 (24.90 %)	114 (30.40 %)	167 (25.38 %)	74 (21.39 %)	351 (25.13 %)	188 (26.07 %)
No	555 (75.10 %)	261 (69.60 %)	488 (74.16 %)	272 (78.61 %)	1,043 (74.66 %)	533 (7.93 %)
Antihypertensive treatment:						
Yes	148 (20.03 %)	88 (23.47 %)	191 (29.03 %)	92 (26.59 %)	339 (24.27 %)	180 (24.97 %)
No	581 (78.62%)	281 (74.93 %)	467 (70.97 %)	254 (73.41 %)	1,048 (75.02 %)	535 (74.20 %)
Systolic Blood Pressure (mmHg)	126.36 ± 13.96	126.79 ± 13.46	122.25 ± 21.58	119.78 ± 19.26	124.36 ± 18.19	123.36 ± 16.91
Total Cholesterol (mg/dL)	190.33 ± 48.84	186.91 ± 44.50	177.68 ± 44.26	178.51 ± 43.98	184.31 ± 47.12	182.87 ± 44.41
HDL (mg/dL)	44.31 ± 14.16	44.91 ± 14.47	48.91 ± 19.41	47.00 ± 18.13	46.50 ± 17.01	45.91 ± 16.35
Other ASCVD risk factors						
Dyslipidemia:						
Yes	554 (74.97 %)	266 (70.93 %)	195 (29.64 %)	94 (27.17 %)	749 (53.61 %)	360 (49.93 %)
No	105 (14.21 %)	76 (20.27 %)	416 (63.22 %)	228 (65.90 %)	521 (37.29 %)	304 (42.16 %)
Missing: n (%)	80 (10.83 %)	33 (8.80 %)	47 (7.14 %)	24 (6.94 %)	127 (9.09 %)	57 (7.91 %)
Statin Medication:						
Yes	162 (21.92 %)	81 (21.60 %)	44 (6.69 %)	16 (4.62 %)	206 (14.75 %)	97 (13.45 %)
No	577 (78.08 %)	294 (78.40 %)	614 (93.31 %)	330 (95.38 %)	1,191 (85.25 %)	624 (86.55 %)

Table 1 Baseline characteristic of MWCCS participants in 2005 in training and testing data sets (continued)

Characteristics	MACS (n = 1,114)		WIHS (n = 1,004)		MWCCS (n = 2,118)	
	Training data set (n = 739)	Testing data set (n = 375)	Training data set (n = 658)	Testing data set (n = 346)	Training data set (n = 1,397)	Testing data set (n = 721)
Self-perceived menopausal status:						
Yes	NA	NA	222 (33.74 %)	120 (34.68 %)	NA	NA
No	NA	NA	414 (62.92 %)	216 (62.43 %)	NA	NA
BMI (kg/m^2)	25.56 $\pm$ 4.35	25.66 $\pm$ 4.27	28.13 $\pm$ 7.47	27.98 $\pm$ 6.78	26.77 $\pm$ 6.16	26.78 $\pm$ 5.74
Waist circumference (cm)	91.63 $\pm$ 11.40	91.35 $\pm$ 11.62	91.20 $\pm$ 15.50	91.31 $\pm$ 14.31	91.43 $\pm$ 13.46	91.33 $\pm$ 12.94
Missing: n (%)	107 (14.48 %)	41 (10.93 %)	106 (16.11 %)	53 (15.32 %)	213 (15.25 %)	94 (13.04 %)
LDL (mg/dL)	112.11 $\pm$ 38.23	110.04 $\pm$ 36.25	98.85 $\pm$ 36.03	101.38 $\pm$ 36.23	105.05 $\pm$ 37.65	105.42 $\pm$ 36.47
Missing: n (%)	176 (23.82 %)	87 (23.20 %)	17 (2.58 %)	17 (4.91 %)	193 (13.82 %)	104 (14.42 %)
Triglyceride (mg/dL)	138.00 (129.00)	138.00 (133.00)	127.50 (100.25)	126.00 (93.75)	133.00 (110.00)	132.00 (109.50)
Missing: n (%)	174 (23.55 %)	86 (22.93 %)	2 (0.30 %)	4 (1.16 %)	176 (12.60 %)	90 (12.48 %)
PLHIV-specific risk factors						
CD4 (cell/mm^3)	505.00 (360.00)	499.00 (369.50)	412.00 (360.00)	409.00 (380.50)	464.50 (371.50)	459.50 (387.25)
CD8 (cell/mm^3)	861.00 (524.00)	863.50 (575.50)	803.00 (577.50)	763.00 (484.75)	840.00 (549.25)	808.00 (513.50)
Viral Load (count/mm^3)	50.00 (6,960)	50.00 (4,790)	98.00 (5,970.00)	80.00 (3,495)	80.00 (6,450.00)	80 (3,850.00)
Duration on PI (year)	1.86 (5.87)	2.44 (6.34)	3.75 (7.00)	3.75 (7.00)	2.75 (6.58)	3.00 (6.75)
Duration on NNRTI (year)	1.00 (3.58)	0.92 (3.19)	1.25 (3.25)	0.75 (3.25)	1.08 (3.33)	0.78 (3.25)
Duration on Abacavir (year)	0.00 (1.07)	0.00 (1.38)	0.00 (1.25)	0.00 (1.25)	0.00 (1.25)	0.00 (1.25)
Inflammatory biomarkers						
CRP (mg/dL)	1.40 (3.10)	1.30 (2.50)	2.05 (4.10)	1.90 (3.90)	1.60 (3.50)	1.60 (2.90)
Missing: n (%)	57 (7.71 %)	27 (7.20 %)	160 (24.32 %)	82 (23.70 %)	217 (15.53%)	109 (15.12 %)

Figures are presented as mean $\pm$ standard deviation or median (interquartile range) for continuous variables, depending on their distribution, and count (percentage) for categorical variables.
Some cells have cumulative percentage slightly higher or lower than 100% from rounding.

Table 2 Newly derived models from the training data sets

Duration	MACS sex-specific model		WIHS sex-specific model		MWCCS model	
	Predictors	Coefficients	Predictors	Coefficients	Predictors	Coefficients
5-year duration	LDL	-4.9044×10^{-3}	TC	-4.1524×10^{-3}	Log(AGE)	4.1387×10^{-1}
	CD4	-2.5283×10^{-6}	HDL	-3.4584×10^{-4}	TC	-5.9622×10^{-4}
	AGE*WC	1.3475×10^{-4}	CD4	-1.1175×10^{-3}	Log(HDL)	-9.7884×10^{-2}
	AGE*SBP	1.2895×10^{-4}	Log(AGE)*SMKGRP1	1.8679×10^{-1}	LDL	-1.9982×10^{-3}
	AGE*DM0	2.2432×10^{-3}	Log(AGE)*Log(VL)	9.4339×10^{-3}	Log(CD4)	-2.4443×10^{-1}
	AGE*Sqrt(ABC)	2.6529×10^{-3}	Log(AGE)*CRP	1.8994×10^{-3}	SEX0*SMKGRP1	4.1759×10^{-1}
	AGE*Log(VL)	1.1958×10^{-3}	RACE1*Log(VL)	1.4285×10^{-2}	Log(AGE)*SMKGRP1	2.6438×10^{-2}
	AGE*CRP	4.4308×10^{-4}	RACE1*MENO1	2.0429×10^{-1}	Log(AGE)*Log(VL)	2.9628×10^{-3}
	RACE0*DM1	-2.1927×10^{-1}			Log(AGE)*CRP	3.1966×10^{-3}
	RACE0*DYSLIP1	-1.2240×10^{-1}			RACE1*Log(VL)	3.2144×10^{-2}
	RACE0*LDL	-2.5015×10^{-3}				
	RACE1*Log(VL)	9.5981×10^{-3}				
	RACE0*CRP	5.4763×10^{-3}				
	Baseline survival; $S_0(5)$	0.9877	Baseline survival; $S_0(5)$	0.9373	Baseline survival; $S_0(5)$	0.9528
10-year duration	Log(AGE)	1.3397	TC	-2.2558×10^{-4}	Log(AGE)	1.1507
	Log(HDL)	-1.1094×10^{-1}	Log(HDL)	-1.8415×10^{-1}	Log(HDL)	-1.7394×10^{-1}
	LDL	-8.7132×10^{-3}	Sqrt(CD4)	-3.4115×10^{-2}	LDL	-1.7487×10^{-3}
	CD4	-3.8906×10^{-4}	MENO1	6.4678×10^{-2}	Log(CD4)	-1.9059×10^{-1}
	Log(AGE)*SMKGRP1	1.1012×10^{-1}	Log(AGE)*SMKGRP1	1.7418×10^{-1}	SEX0*SMKGRP1	6.5809×10^{-1}
	Log(AGE)*Log(SBP)	2.9258×10^{-1}	Log(AGE)*SBP	8.6780×10^{-4}	SEX0*HTNRX1	1.7165×10^{-1}
	Log(AGE)*DM0	1.1610×10^{-2}	Log(AGE)*HTNRX1	2.3078×10^{-2}	Log(AGE)*SMKGRP1	3.6563×10^{-2}
	Log(AGE)*STATIN0	4.7247×10^{-2}	Log(AGE)*Sqrt(PI)	6.7098×10^{-4}	Log(AGE)*SBP	6.6744×10^{-4}
	Log(AGE)*Sqrt(PI)	5.6777×10^{-3}	Log(AGE)*Log(VL)	1.5308×10^{-2}	Log(AGE)*Sqrt(PI)	4.8305×10^{-3}
	Log(AGE)*Log(TG)	2.2864×10^{-3}	Log(AGE)*MENO1	3.3940×10^{-2}	Log(AGE)*Log(VL)	1.2417×10^{-2}
	Log(AGE)*Log(VL)	1.9388×10^{-2}	RACE0*SMKGRP0	-2.0910×10^{-1}	Log(AGE)*CRP	1.5154×10^{-3}
	Log(AGE)*CRP	5.3032×10^{-4}	RACE0*HTNRX0	-4.4819×10^{-2}	RACE0*SMKGRP0	-1.5278×10^{-1}
	RACE0*SMKGRP0	-1.4719×10^{-4}	RACE1*Log(VL)	3.5500×10^{-3}	RACE0*HTNRX0	-3.3292×10^{-2}
	RACE0*DM1	-5.8193×10^{-1}	RACE1*MENO1	2.1420×10^{-1}	RACE0*DYSLIP1	-7.1972×10^{-2}
	RACE0*HTNRX0	-2.822×10^{-1}			RACE1*Log(VL)	2.0527×10^{-2}
	RACE0*DYSLIP0	2.5027×10^{-1}				
	RACE0*DYSLIP1	-2.5576×10^{-1}				
	RACE1*STATIN1	-1.8907×10^{-1}				
	RACE0*Exp(ABC)	1.2449×10^{-3}				
	RACE0*CRP	2.4037×10^{-2}				
	Baseline survival; $S_0(10)$	0.9999	Baseline survival; $S_0(10)$	0.8244	Baseline survival; $S_0(10)$	0.9953

To calculate individual risk over 5 or 10 years, multiply the value of that individual predictor by the table accompanying coefficient. Sum the resulting values to yield a quantity "A". Then compute B = exp(A). Consequently, the risk is equal to $1 - S_0(5)^B$ or $1 - S_0(10)^B$ for 5-year and 10-year period of each cohort, respectively. Abbreviation for transformations: square root (Sqrt), natural logarithm (Ln).

Table 3 Number of deaths, cumulative follow-up time, and average follow-up time in testing data sets

Duration	testing data sets	Number of deaths	Cumulative follow-up time (person-year)	Average follow-up time (years)
5-year duration	MACS	14	1,734	4.60
	WIHS	16	1,594	4.60
	MWCCS	30	3,328	4.60
10-year duration	MACS	21	3,246	8.70
	WIHS	49	2,788	8.10
	MWCCS	70	6,034	8.40

Table 4 Model performance of each newly derived model in testing data sets

Duration	Testing data sets	Sex-specific models			MWCCS model			Benchmarking models*
		AUC	NRI_e	NRI_{ne}	AUC	NRI_e	NRI_{ne}	AUC
5-year duration	MACS	0.61 (0.51,0.73)	-0.14 (-0.46,0.18)	0.24 (0.19,0.30)	0.67 (0.54,0.82)	-0.28 (-0.55,-0.00)	0.29 (0.24,0.34)	0.59 (0.45,0.69)
	WIHS	0.64 (0.50,0.78)	-1.00 (-1.00,-0.99)	0.99 (0.98,1.00)	0.64 (0.50,0.77)	-1.00 (-1.00,-0.99)	0.99 (0.98,1.00)	0.60 (0.48,0.73)
	MWCCS	0.63 (0.54,0.73)	-0.61 (-0.83,-0.36)	0.60 (0.56,0.64)	0.64 (0.56,0.75)	-0.67 (-0.86,-0.47)	0.63 (0.59,0.66)	0.56 (0.44,0.67)
10-year duration	MACS	0.66 (0.57,0.76)	0.54 (0.31,0.76)	-0.06 (-0.14,0.01)	0.67 (0.56,0.79)	0.63 (0.44,0.81)	-0.26 (-0.34,-0.18)	0.47 (0.36,0.59)
	WIHS	0.64 (0.54,0.85)	0.71 (0.52,0.87)	-0.58 (-0.65,-0.52)	0.68 (0.60,0.76)	0.79 (0.61,0.93)	-0.62 (-0.69,-0.56)	0.56 (0.49,0.65)
	MWCCS	0.68 (0.57,0.76)	0.68 (0.53,0.79)	-0.29 (-0.35,-0.25)	0.70 (0.64,0.76)	0.75 (0.62,0.87)	-0.42 (-0.47,-0.37)	0.49 (0.42,0.55)

Benchmarking models are the VACS Index for 5-year duration and the PCE for 10-year duration. Abbreviation: area under receiver operating characteristic curve (AUC), net-reclassification of event (NRI_e), net-reclassification of non-event (NRI_{ne})

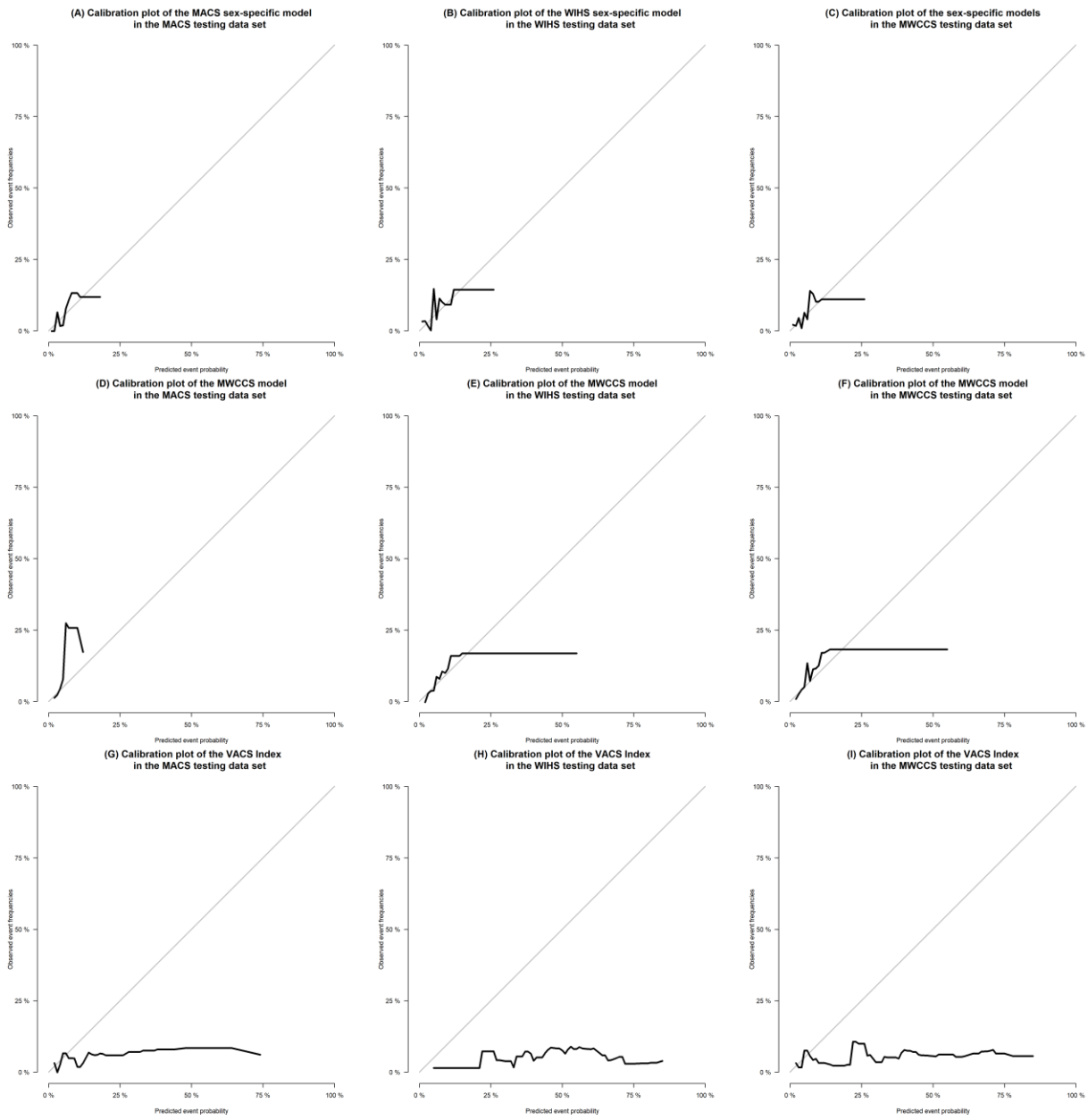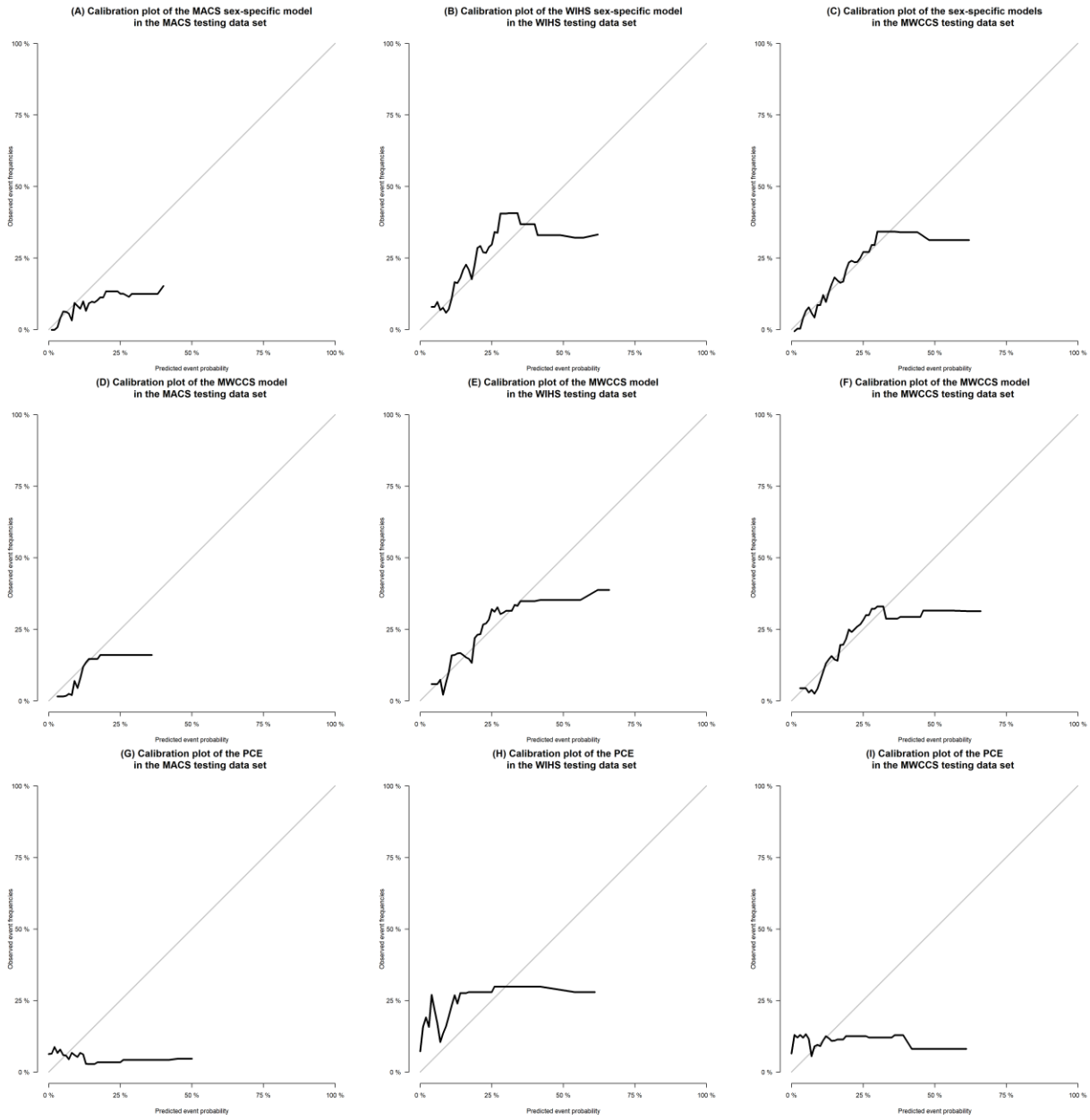**Figure 1 5-year calibration plots in testing data sets**

Figure 2 10-year calibration plots in testing data sets

4.3 Results

We showed the newly derived models in Table 2. The table contains predictors, coefficients, and baseline survival probabilities for both 5- and 10-year periods. Overall, these models contain eight to twenty predictors. Some predictors occur repeatedly in several models including age, race, CD4, and smoking group. Additional information on penalty factors of these models is available in the Appendix (Appendix B supplementary table 1). The accompanying residuals plots suggest that the models respect the proportional hazard assumption (Appendix B supplementary figures 1 to 6). Furthermore, we provided an example of individual risk calculation based on the 10-year MWCCS model in the supplementary text (Appendix B supplementary text).

Table 3 describes the number of deaths and the person-time in each testing data set. At 5 years, there were 14 deaths among the MACS and 16 deaths among the WIHS, resulting in 30 deaths among the MWCCS. Divided by the cumulative person-years, the 5-year crude all-cause mortality rate among the MACS, the WIHS and the MWCCS were 8.07 deaths per 1,000 PLHIV-year, 10.04 deaths per 1,000 PLHIV years and 9.01 deaths per 1,000 PLHIV-year, respectively. At 10 years, there were 21 deaths among the MACS and 49 deaths among the WIHS, adding up to 70 deaths among the MWCCS. With the same calculation, we obtained the 10-year crude all-cause mortality rate of 6.47 deaths per 1,000 PLHIV-year for the MACS, 17.58 deaths per 1,000 PLHIV-years for the WIHS and 11.06 deaths per 1,000 PLHIV-year for the MWCCS. The average follow-up time was 4.6 years for the 5-year period and more than 8 years for the 10-year period. This high average follow-up time indicated that most of the participants had complete follow-up visits over these two time periods.

Table 4 illustrates discrimination property of each model by AUC. For 5-year duration, the newly derived models had point estimations of AUC slightly higher than that of the VACS Index; however, the 95%CI of these figures largely overlapped. For the 10-year MACS testing data set, the 95%CI overlap just slightly between the lower limits of the new models and the upper limits of the PCE. In the 10-year WIHS testing data set, the point estimations of AUC were higher from the new models. However, their 95%CI extended over the corresponding interval from the PCE. For 10-year MWCCS testing data set, both the sex specific models and the MWCCS model had their entire 95%CI of the AUC above that of the PCE. Accompanying ROC curves are provided in the Appendix B (Appendix B supplementary figures 7 and 8).

Figures 1 and 2 are calibration plots over the 5-year duration and the 10-year duration, respectively. The degree of good calibration is reflected by the length of the agreement line, a line that represents the agreement between the observed event frequency over the follow-up period on the y-axis and the predicted probability of events on the x-axis, that situates on the diagonal line of the calibration plot. For 5-year duration, figure 1 illustrates that the new models had at least some part of their plots on the diagonal line (figure 1, plots A to F). On the other hand, the VACS Index aligned its agreement far from the diagonal line (figure 1, plots E to I). Furthermore, the VACS Index had its entire agreement lines on the lower right portion of calibration plot. This pattern indicated that the VACS Index overestimated the risk of death over the 5-year duration in each testing data set. For the 10-year duration, all new models were well aligned to the diagonal line for most part of their agreement line (figure 2, plots A to F). The calibration plots of the PCE were not well aligned with the diagonal line among the MACS and the MWCCS testing data sets (figure 2, plots E and H). Nevertheless, we could observe some trend from the calibration plot of the PCE in the WIHS testing data set that conformed with the

90

slope of the diagonal line (figure 2, plot I). Additionally, the PCE's calibration plots had initial

portion of its plot in the lower left quadrant above the diagonal line reflecting a varying degree of

under estimation of risk in each testing data set.

In terms of reclassification, table 4 includes the net reclassification of events (NRI_e) and

non-event (NRI_{ne}). Over five years, all NRI_e were negative while all NRI_{ne} were positive. This

finding suggested that the risk prediction from the new models, as compared to the risk from the

VACS Index, had negative net proportion of events assigned to the higher risk category and

positive net proportion of non-events assigned to the lower risk category. Over 10 years, all NRI_e

were positive while all NRI_{ne} were negative. This result reflects that the risk from the new

models had positive net proportion of events assigned to the higher risk category and negative

net proportion of non-events assigned to the lower risk category. These net reclassifications

complied with the previous findings from the calibration plots that the VACS Index tended to

overestimate the risk in the 5-year period and the PCE appeared to have some degree of

underestimation of the risk over the 10-year period.

4.4 Discussion

Our analysis supported the utilization of available cardiovascular and related clinical risk

factors to predict all-cause mortality among PLHIV. The elastic net regularization resulted in

prediction models that were guarded against overfitting. Improvement in discrimination was seen

from the sex-specific models and the MWCCS model with the evaluation among the 10-year

MWCCS testing data set and also, to a lesser extent, among the 10-year MACS testing data set.

All calibration plots pointes towards better calibration from all newly derived model as

compared to the VACS Index at 5 years and the PCE at 10 years. The NRI_e and the NRI_{ne} also indicated that the new models reclassify participants in a reasonable direction as compared to the misestimation observed from the benchmarking models.

The comparable discrimination between the new models and the VACS Index could be explained by the shared predictors between the models which are Age, CD4 counts, and viral load. These variables are all important predictors of PLHIV survival and appears in both the new models and the VACS Index [10–12]. For calibration, the overestimation of mortality risk from the VACS Index could be explained by the different characteristics of the participants in the Veterans study [10–12] and the MWCCS. Moreover, we did not recalibrate coefficients in the VACS Index to represent comparison to the routine use in clinical setting.

Over a 10-year period, the improved performance of the new models as compared to the PCE could be due to two reasons. From the mechanistic perspective, the new models had an advantage of including more PLHIV predictors than the PCE. From modeling standpoint, the implementation of elastic net regularization further finetuned the model[9]. However, one might argue that this improvement is expected from the fact that PCE itself does not directly predict all-cause mortality[13]. Nevertheless, our rationale of this comparison is to make comparison to the standard prediction model in clinical care and that the number of ASCVD death among PLHIV has been increasing over the years.

This study had a few limitations. First, we had a low number of events of interest. This limited number of events was further affected by splitting of data sets into training and testing data sets. Additionally, the number of events in each data sets grew fewer from the stratification by time duration and sex. The few events could potentially lead to overfitting. Nevertheless, we at least counteracted this shortcoming with the use of regularization. Second, we were unable to

evaluate some variables due to high missingness such as IL-6. Third, we encountered modeling limitations stemmed from the glmnet package for the cox model. To our knowledge, the current function of glmnet for cox model does not respect hierarchical structure. Unsurprisingly, we obtained prediction models with several interaction terms, but without the main effects. While this modeling approach might be forgiving for prediction, this model specification raises a concern if we are to apply our models to the causal inference context. Fourth, our assumptions on handling of missing data could be challenged. Besides the single imputation that we did, multiple imputation might be considered in this context as well. Despite these limitations, our study had several strengths. First, we utilized data from long-standing cohort studies which is an appropriate study design for building prediction models. Second, we incorporated a modern statistical approach to improve predictive performance of the new models. Third, we offered a range of predictive performance measures from different angles not only in terms of absolute quantitative and qualitative measures, such as the AUC and the calibration plots, but also relative to prior models, which were the NRI_e and NRI_{ne}.

4.5 Public health implications

Over 10-year period, the sex-specific models and the MWCCS model illustrated satisfactory performance for all three measures in the MWCCS testing data set. Therefore, these models should be extended to external validation to test its performance in other settings.

4.6 Conclusion

Both the 10-year sex-specific models and the 10-year MWCCS model showed better discrimination than the PCE in the MWCCS testing data set. These models also possessed good calibration and appropriate risk reclassification. Therefore, these models should be explored in other settings to evaluate their external validity.

4.7 Reference

1   Merson MH. **The HIV-AIDS pandemic at 25 - the global response**. *N Engl J Med* 2006; **354**:6–8.

2   Hogg R, Lima V, Sterne JAC, Grabar S, Battegay M, Bonarek M, *et al.* **Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies**. *Lancet* 2008; **372**:293–299.

3   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**:6–13.

4   Bosh KA, Johnson AS, Hernandez AL, Prejean J, Taylor J, Wingard R, *et al.* **Vital signs: deaths among persons with diagnosed HIV infection, United States, 2010–2018**. *MMWR* 2020; **69**:1717–1724.

5   Feinstein MJ, Bahiru E, Achenbach C, Longenecker CT, Hsue P, So-Armah K, *et al.* **Patterns of cardiovascular mortality for HIV-infected adults in the United States: 1999 to 2013**. *Am J Cardiol* 2016; **117**:214–220.

6   Adih WK, Selik RM, Hall HI, Babu AS, Song R. **Associations and trends in cause-specific rates of death among persons reported with HIV infection, 23 U.S. jurisdictions, through 2011**. *Open AIDS J* 2016; **10**:144–157.

7   Wiemken TL, Kelley RR. **Machine learning in epidemiology and health outcomes research**. *Annu Rev Public Health* 2020; **41**:21–36.

8   Steyerberg EW. Overfitting and Optimism in Prediction Models. In: *Clinical Prediction*

Models 2E. Gail M, Samet JM, Singer B (editors). . Cham: Springer; 2019. pp. 95–102.

9   Steyerberg EW. Penalized Estimation. In: *Clinical Prediction Models 2E*. Gail M, Samet JM, Singer B (editors). . Cham: Springer; 2019. pp. 251–259.

10   Tate JP, Justice AC, Hughes MD, Bonnet F, Reiss P, Mocroft A, *et al.* **An internationally generalizable risk index for mortality after one year of antiretroviral therapy**. *AIDS* 2013; **27**:563–572.

11   Justice A, Modur S, Tate J, Althoff K, Jacobson L, 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**. *J Acquir Immune Defic Syndr* 2013; **62**:149–163.

12   VACS Index Calculator. 2021.https://vacs.med.yale.edu/calculator/IC (accessed 8 Mar2021).

13   Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, *et al.* **2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines**. *J Am Coll Cardiol* 2014; **63**:2935–2959.

14   Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR. **The Multicenter AIDS Cohort Study: rationale , organization, and selected characteristics of the participants**. *Am J Epidemiol* 1987; **126**:310–318.

15   Detels R, Jacobson L, Margolick J, Martinez-Maza O, Muñoz A, Phair J, *et al.* **The Multicenter AIDS Cohort Study, 1983 to ...** *Public Health* 2012; **126**:196–198.

16   Barkan SE, Melnick SL, Preston-martin S, Weber K, Kalish LA, Miotti P, *et al.* **The**

Women' s Interagency HIV Study. *Epidemiology* 1998; **9**:117–125.

17   Bacon MC, Von Wyl V, Alden C, Sharp G, Robison E, Hessol N, *et al.* **The Women's Interagency HIV Study: an observational cohort brings clinical sciences to the bench**. *Clin Diagn Lab Immunol* 2005; **12**:1013–1019.

18   Adimora AA, Ramirez C, Benning L, Greenblatt RM, Kempf MC, Tien PC, *et al.* **Cohort profile: the Women's Interagency HIV Study (WIHS)**. *Int J Epidemiol* 2018; **47**:393-394I.

19   R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

20   Jerome Friedman, Trevor Hastie, Robert Tibshirani (2010). **Regularization Paths for Generalized Linear Models via Coordinate Descent**. Journal of Statistical Software, 33(1), 1-22. URL https://www.jstatsoft.org/v33/i01/.

21   Hong Ooi (2021). glmnetUtils: Utilities for 'Glmnet'. R package version 1.1.8. https://CRAN.R-project.org/package=glmnetUtils

22   Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning 2E*. New York: Springer series in statistics; 2009. http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf

23   Patrick J. Heagerty and packaging by Paramita Saha-Chaudhuri (2013). survivalROC: Time-dependent ROC curve estimation from censored survival data. R package version 1.0.3. https://CRAN.R-project.org/package=survivalROC

24   Ulla B. Mogensen, Hemant Ishwaran, Thomas A. Gerds (2012). **Evaluating Random**

Forests for Survival Analysis Using Prediction Error Curves. Journal of Statistical

Software, 50(11), 1-23. URL https://www.jstatsoft.org/v50/i11.

25   Eisuke Inoue (2018). nricens: NRI for Risk Prediction Models with Time to Event and

Binary Response Data. R package version 1.6. https://CRAN.R-

project.org/package=nricens

Manuscript #3

Prediction model for non-HIV mortality from cardiovascular risk factors

among people living with HIV

Abstract

Background: A high proportion of people living with HIV (PLHIV) had cardiovascular causes as the significant contributor; therefore, cardiovascular risk factors should be predictive for non-HIV mortality among PLHIV.

Methods: We fitted the Fine and Gray model with boosting algorithm separately by sex and for the whole MWCCS (MACS/WIHS Combined Cohort Study) in 2005 to predict non-HIV mortality over five and ten years. We set out to compare the new model against the Veterans Aging Cohort Study (VACS) Index and the pooled cohort equations (PCE) for discrimination, calibration, and net reclassification.

Results: We included 1,114 male and 1,004 female participants in this analysis. The proportion of non-HIV death to total death was 30% for 5-year period and 45% for 10-year period. Only 10-year WIHS sex-specific model and 10-year MWCCS model satisfied the sub-distributional hazard assumption based on residual plots. None of the two new models demonstrated satisfactory discrimination. Calibration plots suggested that the new model had slightly better calibration. The net reclassification indicated that the new models classified participants to higher and lower risk group in different proportion as compared to the PCE.

Conclusion: The new model appears to slightly improved calibration and had different risk reclassification than the PCE. However, their discrimination property was not satisfactory.

5.1 Background

One of the most ground-breaking interventions in the field of HIV is combined antiretroviral therapy (cART). The triple-drug regimen effectively halts the progression of the disease by intervening at several checkpoints in the viral life cycle. Consequently, HIV-related deaths among PLHIV decreased significantly, and PLHIV survived longer[1].

The prolonged survival led PLHIV to experience non-communicable diseases at a younger age compared to the age-matched non-HIV population. The increase in non-communicable disease burden among PLHIV later translated to the rise in non-HIV mortality among PLHIV[2,3]. Therefore, quantifying the risk of death from non-HIV causes could benefit PLHIV counseling in the post-cART era.

The analysis of HIV surveillance in the United States (US) revealed that non-HIV mortality contributed more to the total death among PLHIV compared to HIV-related causes in these recent years[4]. A closer look at the data showed that the cardiovascular death is the highest attributable cause of death among non-HIV mortality[5]. Consequently, it is reasonable that cardiovascular risk factors should be predictive of non-HIV mortality among PLHIV.

While no established prediction model targets non-HIV mortality specifically, mortality prediction among PLHIV is achievable by the Veteran Aging Cohort Study (VACS) Index. The VACS Index incorporates biologically relevant covariates among PLHIV that yields a 5-year mortality risk [6,7]. Since most of the non-HIV death is attributable to cardiovascular cause, a prediction from the cardiovascular realm should be considered. For the general population in the US, the pooled cohort equations (PCE), sex-race-specific equations from five large cohorts in the US, predict the 10-year risk of atherosclerotic cardiovascular disease (ASCVD) [8].

VACS Index and PCE share two common drawbacks in prediction from the modeling perspective. First, both survival models are not based on a specific regression model in competing risk settings. Survival analysis in the presence of competing risk needs a more precise model than the standard cox proportional hazard model[9–11]. Second, both models do not incorporate more modern techniques to avoid overfitting. Several machine learning tools, such as ensemble method, can minimize overfitting[12].

Therefore, we developed and evaluated a prediction model of non-HIV mortality among PLHIV based on cardiovascular risk factors utilizing the survival model for competing risk with an ensemble technique to deal with overfitting. We compare our newly derived model to the VACS index for 5-year prediction and the PCE for 10-year prediction.

5.2 Methods

5.2.1 Source of Data and study population

We use data from the Multicenter AIDS Cohort Study (MACS) and Women's Interagency HIV Study (WIHS) Combined Cohort Study (MWCCS). Details on study design were available elsewhere[13–17]. Briefly, MACS and WIHS are multicenter longitudinal cohort studies on HIV among gay men and high-risk women in the US, respectively. Both studies had comparable biannual visits which included interviews, physical examination, blood test, and other clinical workups.

In this analysis, the baseline study population was all MWCCS participants in 2005. We select this time point since the change in mortality pattern occurred during the early 2000s and also because of the availability of the variables for risk calculation of the benchmarking models.

This process resulted in 1,114 MACS participants and 1,004 WIHS participants included in this study. The baseline characteristic, as randomly split into training and testing data sets, is showed in Table 1. Apart from waist circumference (WC), low-density lipoprotein cholesterol (LDL), triglyceride (TG), and C-reactive protein (CRP), all other variables have missing values less than 10%. This research protocol was approved by the Institutional Review Board in University of California, Los Angeles (IRB# 20-001446) and the MWCCS Executive Committee (Project# X20051).

5.2.2 Predictors

Twenty-four potential predictor variables from the MWCCS were chosen for this analysis. We drew on the literature on predictors for ASCVD, both in traditional practice and PLHIV-specific situation, for the selection of these variables[2,18,19]. Each participant had covariate values from nearest visit within 2004 to 2006 window. We carried out single imputation of missing data by median (continuous variables) or mode (categorical variables). The predictors were classified into the four categories, as listed below. Each predictor was presented with their study abbreviation together with given numerical codes of each level or unit of measurement for categorical variables and continuous variables, respectively.

5.2.2.1 Traditional ASCVD risk factors

Traditional ASCVD risk factors included the following variables: age (AGE, years), sex (SEX; male = 1, female =0), race (RACE; African American = 1, White and others = 0), smoking status (SMKGRP; current smoker = 1, non or past smoker = 0), diabetes mellitus (DM; yes = 1, no = 0), hypertension (HTN; yes = 1, no = 0), antihypertensive treatment (HTNRX;

received treatment = 1, not received treatment = 0), systolic blood pressure (SBP; mmHg), total

cholesterol (TC; mg/dL), and high-density lipoprotein cholesterol (HDL; mg/dL).

5.2.2.2 Other ASCVD risk factors

Aside traditional ASCVD predictors, this domain encompassed other potential predictors

for ASCVD: dyslipidemia (DYSLIP; yes = 1, no = 0), statin medication use (STATIN; yes = 1,

no = 0), body mass index (BMI; kg/m^2), waist circumference (WC; cm), low-density lipoprotein

cholesterol (LDL; mg/dL) and triglyceride level (TG; mg/dL). For the WIHS, self-perceived

menopausal status (MENO; yes = 1, no = 0) was also included.

5.2.2.3 PLHIV-specific risk factors

HIV-related blood tests and cART exposure were incorporated into our model. The blood

tests were CD4+ cell count (CD4; cell/mm3), CD8+ cell count (CD8; cell/mm3), and HIV RNA

viral load (VL; copies/ ml). Cumulative years of protease inhibitor (PI; years), non-

nucleotide/nucleoside reverse transcriptase inhibitor (NNRTI; years), and Abacavir (ABC; years)

were the cART exposure variables.

5.2.2.4 Inflammatory biomarkers

On top of PLHIV-specific risk factors that are part of the disease-causing process, other

blood tests that might reflect inflammatory response but not be specific to HIV pathophysiology,

such as c-reactive protein (CRP; mg/dL) and interleukin-6 (IL-6; pg/dL), were considered.

Nevertheless, we later excluded IL-6 due to a high proportion of missing values. CRP was

calculated as a weighted mean from each available CRP result among WIHS participants.

5.2.3 Outcome

We included non-HIV deaths after 2005 in this analysis. Non-HIV deaths were determined based on the ICD code from the underlying cause of death. We excluded the HIV-related death based on ICD-10 of B21-24 or ICD-9 of 42. Furthermore, we excluded deaths due to external causes, which encompassed accidents, intentional self-harm, assaults (ICD-10 code V01-Y89 or corresponding ICD-9 code starting with "E") or use of a psychoactive substance (ICD-10 code F11-F16, F18-F19 or corresponding ICD-9 code 29XXX & 30XXX with Xs as placeholders for numbers in the ICD-9 code). For both cohorts, deaths were ascertained from the National Death Index and death certificates. Person time was enumerated in person-years. We assumed that loss-to-follow-up occurred between the last-know visit and the subsequent missed visit. For participants who were lost to follow-up during 2006-2015 and later returned to MWCCS, we assigned full person time of this study period.

5.2.4 Statistical analysis: model development and evaluation

We randomly split the original data into training (2/3 of the data) and testing (1/3 of the data) data sets. In the training data set, models for two different prediction periods were fitted. We used data from 2006 to 2010 for 5-year projection and data from 2006 to 2015 for 10-year projection. Within each projection, cohort specific models for MACS, WIHS, and MWCCS were derived. These strategies led to six new models: 5-year MACS sex-specific model, 5-year WIHS sex-specific model, 5-year MWCCS model, 10-year MACS sex-specific model, 10-year WIHS sex-specific model and 10-year MWCCS model. We would later evaluate these models along with the benchmarking models in the testing data set. All the statistical analyses were carried out in R version 4.1.1[20].

Each variable in the training data set underwent the following process during the modeling building step. Each categorical variable was separately coded for each level such that SEX1 = 1 was for males and SEX0=1 was for females. Then, we created two-way interaction terms between age, sex, and race with all other variables. Consequently, non-linearity of each continuous predictor was explored using a generalized additive model controlling for age, sex, and race. Finally, the following transformation of each continuous variable was considered: cubic, quadratic, square root, logarithmic, and exponential transformation.

We chose the sub-distribution proportional hazard model rather than the cause-specific proportional hazard model for our risk prediction model in the context of competing risk based on recent literature on this topic[9–11]. We processed the predictor variables through boosting algorithm via cross-validation from the Coxboost package[21]. Boosting is an ensemble method that trains models on a subset of data in sequential order. Predictive errors from each round improved the fitting in the next round. The final model was an average across each model from the whole process[12]. The Coxboost package utilizes an offset-based boosting approach. Each boosting step carries over to the next step as penalized partial likelihood estimation one covariate at a time. The process leads to many estimated zero coefficients; therefore, the final model has the same variable selection as the lasso regularization. The cross-validation ensures that the model has the number of boosting steps, the tuning parameter of this approach, that leads to the lowest partial loglikelihood[21]. We calculated the predicted probability of non-HIV death over the follow-up period of each model. The sub-distributional proportional hazard assumption was assessed by Schoenfeld residual plots[22].

Predictive performance included discrimination, calibration, and net reclassification. The discrimination measure was the area under the receiver operating characteristic curve (AUC)

using timeROC package[23]. The evaluation of calibration was based on calibration plot by the calPlot function in pec package[24]. Both the net reclassification of events (NRI_e) and the net reclassification of non-events (NRI_{ne}) of each new model with the benchmarking models using the 7.5% and 20% cut-off points via nricens package were calculated[25]. These statistical evaluations have taken competing risk into their assessment. 95% confidence interval (CI) for each measure was obtained from bootstrapping.

5.2.5 Benchmarking models

We initially considered to include two benchmarking models which were the VACS Index and the PCE. However, new models from 5-year training data sets violated the sub-distributional hazard assumption; therefore, we did not further pursue comparison for 5-year period. For 10-year period, the PCE is a set of sex-race specific equations derived from pooled cohorts of the general US population[8]. It yields a 10-year risk of composite ASCVD outcomes: coronary death, non-fatal coronary heart disease, fatal stroke and non-fatal stroke[8]. The predicted 10-year risk was calculated based on the publicly available details on the PCE from the 2013 American College of Cardiology/ American Heart Association (ACC/AHA) report[8]. Any individual predictor value out of the reference range of each risk model was truncated to the upper or lower limits of that model.

Table 1 Baseline characteristics in 2005 of MACS, WIHS and MWCCS participants in training and testing data sets

Characteristics	MACS (n = 1,114)		WIHS (n = 1,004)		MWCCS (n = 2,118)	
	Training data set (n = 739)	Testing data set (n = 375)	Training data set (n = 658)	Testing data set (n = 346)	Training data set (n = 1,397)	Testing data set (n = 721)
Traditional ASCVD risk factors						
Age (year)	46.41 ± 8.93	46.60 ± 9.24	45.78 ± 8.13	46.06 ± 7.27	46.12 ± 8.56	46.34 ± 8.36
Race:						
Black	238 (32.21 %)	108 (28.80 %)	321 (48.78 %)	168 (48.55 %)	559 (40.01 %)	276 (38.28 %)
White	501 (67.79 %)	267 (71.20 %)	337 (51.22 %)	178 (51.45 %)	838 (59.99 %)	445 (61.72 %)
Smoking status:						
Current smoker	273 (36.94 %)	117 (31.20 %)	291 (44.22 %)	147 (42.49 %)	564 (40.37 %)	264 (36.62 %)
Non/ past smoker	451 (61.03 %)	249 (66.40 %)	360 (54.71 %)	194 (56.07 %)	811 (58.05 %)	443 (61.44 %)
Diabetes mellitus:						
Yes	162 (21.92 %)	84 (22.40 %)	85 (12.92 %)	43 (12.43 %)	247 (17.68 %)	127 (17.61 %)
No	577 (78.08 %)	291 (77.60 %)	573 (87.08 %)	303 (87.57 %)	1,150 (82.32 %)	594 (82.39 %)
Hypertension:						
Yes	184 (24.90 %)	114 (30.40 %)	167 (25.38 %)	74 (21.39 %)	351 (25.13 %)	188 (26.07 %)
No	555 (75.10 %)	261 (69.60 %)	488 (74.16 %)	272 (78.61 %)	1,043 (74.66 %)	533 (7.93 %)
Antihypertensive treatment:						
Yes	148 (20.03 %)	88 (23.47 %)	191 (29.03 %)	92 (26.59 %)	339 (24.27 %)	180 (24.97 %)
No	581 (78.62%)	281 (74.93 %)	467 (70.97 %)	254 (73.41 %)	1,048 (75.02 %)	535 (74.20 %)
Systolic Blood Pressure (mmHg)	126.36 ± 13.96	126.79 ± 13.46	122.25 ± 21.58	119.78 ± 19.26	124.36 ± 18.19	123.36 ± 16.91
Total Cholesterol (mg/dL)	190.33 ± 48.84	186.91 ± 44.50	177.68 ± 44.26	178.51 ± 43.98	184.31 ± 47.12	182.87 ± 44.41
HDL (mg/dL)	44.31 ± 14.16	44.91 ± 14.47	48.91 ± 19.41	47.00 ± 18.13	46.50 ± 17.01	45.91 ± 16.35
Other ASCVD risk factors						
Dyslipidemia:						
Yes	554 (74.97 %)	266 (70.93 %)	195 (29.64 %)	94 (27.17 %)	749 (53.61 %)	360 (49.93 %)
No	105 (14.21 %)	76 (20.27 %)	416 (63.22 %)	228 (65.90 %)	521 (37.29 %)	304 (42.16 %)
Missing: n (%)	80 (10.83 %)	33 (8.80 %)	47 (7.14 %)	24 (6.94 %)	127 (9.09 %)	57 (7.91 %)
Statin Medication:						
Yes	162 (21.92 %)	81 (21.60 %)	44 (6.69 %)	16 (4.62 %)	206 (14.75 %)	97 (13.45 %)
No	577 (78.08 %)	294 (78.40 %)	614 (93.31 %)	330 (95.38 %)	1,191 (85.25 %)	624 (86.55 %)

Table 1 Baseline characteristic of MWCCS participants in 2005 in training and testing data sets (continued)

Characteristics	MACS (n = 1,114)		WIHS (n = 1,004)		MWCCS (n = 2,118)	
	Training data set (n = 739)	Testing data set (n = 375)	Training data set (n = 658)	Testing data set (n = 346)	Training data set (n = 1,397)	Testing data set (n = 721)
Self-perceived menopausal status:						
Yes	NA	NA	222 (33.74 %)	120 (34.68 %)	NA	NA
No	NA	NA	414 (62.92 %)	216 (62.43 %)	NA	NA
BMI (kg/m^2)	25.56 $\pm$ 4.35	25.66 $\pm$ 4.27	28.13 $\pm$ 7.47	27.98 $\pm$ 6.78	26.77 $\pm$ 6.16	26.78 $\pm$ 5.74
Waist circumference (cm)	91.63 $\pm$ 11.40	91.35 $\pm$ 11.62	91.20 $\pm$ 15.50	91.31 $\pm$ 14.31	91.43 $\pm$ 13.46	91.33 $\pm$ 12.94
Missing: n (%)	107 (14.48 %)	41 (10.93 %)	106 (16.11 %)	53 (15.32 %)	213 (15.25 %)	94 (13.04 %)
LDL (mg/dL)	112.11 $\pm$ 38.23	110.04 $\pm$ 36.25	98.85 $\pm$ 36.03	101.38 $\pm$ 36.23	105.05 $\pm$ 37.65	105.42 $\pm$ 36.47
Missing: n (%)	176 (23.82 %)	87 (23.20 %)	17 (2.58 %)	17 (4.91 %)	193 (13.82 %)	104 (14.42 %)
Triglyceride (mg/dL)	138.00 (129.00)	138.00 (133.00)	127.50 (100.25)	126.00 (93.75)	133.00 (110.00)	132.00 (109.50)
Missing: n (%)	174 (23.55 %)	86 (22.93 %)	2 (0.30 %)	4 (1.16 %)	176 (12.60 %)	90 (12.48 %)
PLHIV-specific risk factors						
CD4 (cell/mm^3)	505.00 (360.00)	499.00 (369.50)	412.00 (360.00)	409.00 (380.50)	464.50 (371.50)	459.50 (387.25)
CD8 (cell/mm^3)	861.00 (524.00)	863.50 (575.50)	803.00 (577.50)	763.00 (484.75)	840.00 (549.25)	808.00 (513.50)
Viral Load (count/mm^3)	50.00 (6,960)	50.00 (4,790)	98.00 (5,970.00)	80.00 (3,495)	80.00 (6,450.00)	80 (3,850.00)
Duration on PI (year)	1.86 (5.87)	2.44 (6.34)	3.75 (7.00)	3.75 (7.00)	2.75 (6.58)	3.00 (6.75)
Duration on NNRTI (year)	1.00 (3.58)	0.92 (3.19)	1.25 (3.25)	0.75 (3.25)	1.08 (3.33)	0.78 (3.25)
Duration on Abacavir (year)	0.00 (1.07)	0.00 (1.38)	0.00 (1.25)	0.00 (1.25)	0.00 (1.25)	0.00 (1.25)
Inflammatory biomarkers						
CRP (mg/dL)	1.40 (3.10)	1.30 (2.50)	2.05 (4.10)	1.90 (3.90)	1.60 (3.50)	1.60 (2.90)
Missing: n (%)	57 (7.71 %)	27 (7.20 %)	160 (24.32 %)	82 (23.70 %)	217 (15.53%)	109 (15.12 %)

Figures are presented as mean $\pm$ standard deviation or median (interquartile range) for continuous variables, depending on their distribution, and count (percentage) for categorical variables.
Some cells have cumulative percentage slightly higher or lower than 100% from rounding.

Table 2 Newly derived models from the training data sets

Duration	MACS sex-specific model		WIHS sex-specific model		MWCCS model	
	Predictors	Coefficients	Predictors	Coefficients	Predictors	Coefficients
10-year duration	NA	NA	Log(AGE)	1.3873×10^{-2}	Log(AGE)	5.2952×10^{-2}
			CD4	-48.6561	CD4	-39.6429
			Log(CD8)	-1.4022×10^{-1}	SEX0*SMKGRP1	1.0863×10^{-1}
			Log(AGE)*SMKGRP1	8.0056×10^{-1}	SEX0*DM1	3.9015×10^{-2}
			Log(AGE)*DM1	2.9891×10^{-1}	SEX0*HTNRX1	1.2650×10^{-2}
			Log(AGE)*PI	9.2917×10^{-1}	SEX0*Sqrt(VL)	16.0232
			Log(AGE)*NNRTI	-2.6808×10^{-1}	Log(AGE)*SMKGRP1	1.9417×10^{-1}
			Log(AGE)*Log(VL)	1.3079	RACE0*HTNRX0	-1.3320×10^{-1}
			RACE0*SMKGRP0	-5.8194×10^{-2}	RACE0*DYSLIP1	-1.0068×10^{-2}
			RACE1*DM1	1.6715×10^{-2}		
			RACE0*HTNRX0	-7.4217×10^{-2}		
			RACE0*HTNRX1	6.1552×10^{-2}		
			RACE0*PI	6.2252×10^{-2}		
			RACE1*MENO1	7.0326×10^{-2}		
			$S_0(10)$	0.9008	$S_0(10)$	0.9997

To calculate individual risk over 5 or 10 years, multiply the value of that individual predictor by the table accompanying coefficient. Sum the resulting values to yield a quantity "A". Then compute B = exp(A). Consequently, the risk is equal to $1 - S_0(5)^B$ or $1 - S_0(10)^B$ for 5-year and 10-year period of each cohort, respectively. Abbreviation for transformations: square root (Sqrt), natural logarithm (Ln).

Table 3 Number of non-HIV deaths, total number of deaths, and average follow-up time in testing data sets
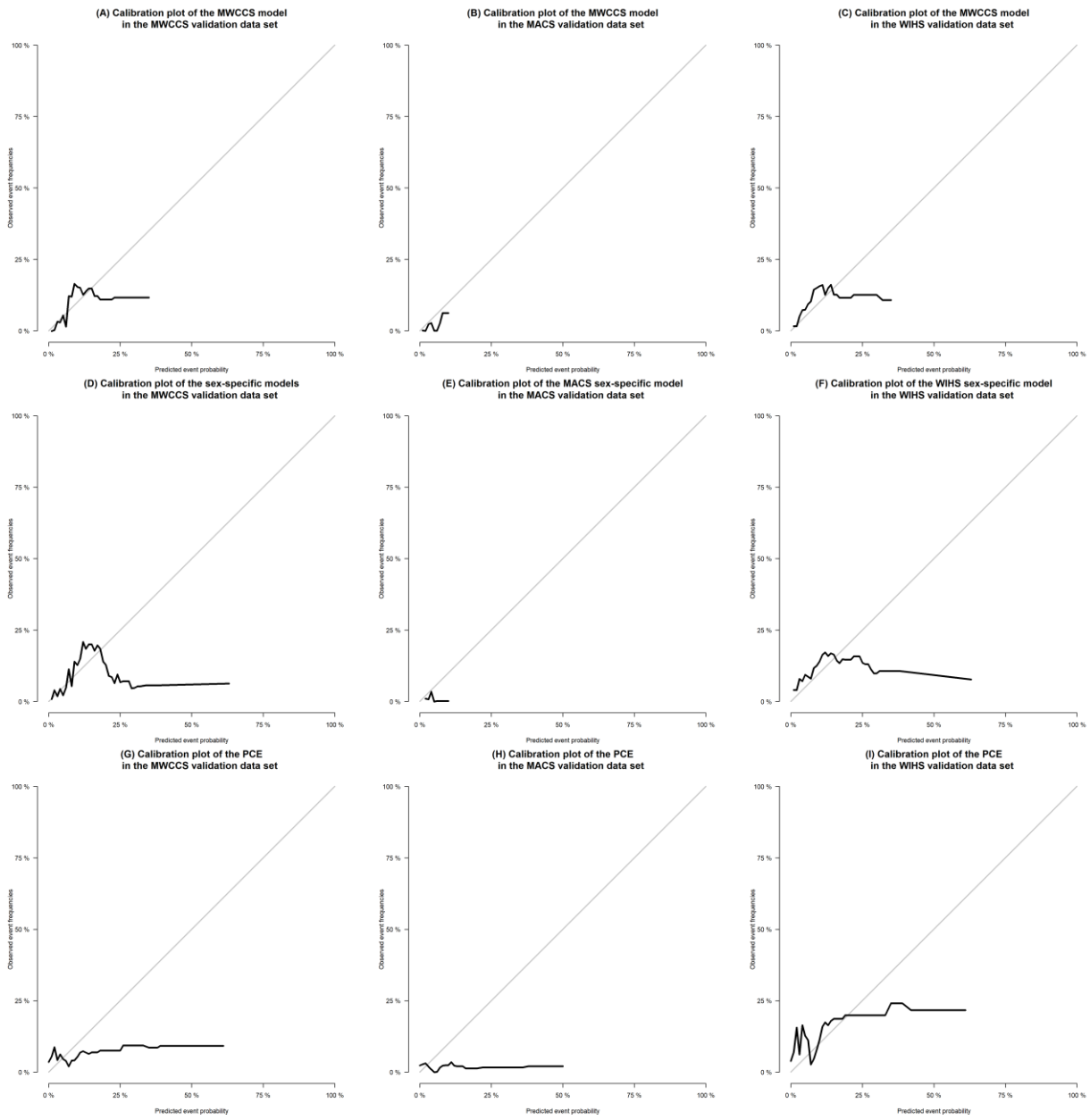
Duration	Testing data sets	Number of non-HIV deaths	Total number of deaths	Average follow-up time (years)
5-year duration	MACS	3	14	4.63
	WIHS	6	16	4.61
	MWCCS	9	30	4.65
10-year duration	MACS	6	21	8.66
	WIHS	26	49	8.10
	MWCCS	32	70	8.46

Table 4 Model performance of each newly derived model in testing data sets

Duration	Testing data sets	Sex-specific models			MWCCS model			PCE
		AUC	NRI$_e$	NRI$_{ne}$	AUC	NRI$_e$	NRI$_{ne}$	AUC
10-year duration	MACS	NA	NA	NA	0.57 (0.22,1.00)	-0.19 (-0.56,0.00)	0.57 (0.22,1.00)	0.44 (0.08,0.93)
	WIHS	0.38 (0.19,0.58)	0.29 (-0.02,0.60)	-0.23 (-0.33,-0.13)	0.39 (0.21,0.61)	0.35 (0.06,0.65)	0.39 (0.21,0.61)	0.62 (0.45,0.79)
	MWCCS	NA	NA	NA	0.54 (0.39,0.70)	0.30 (0.03,0.59)	0.09 (0.04,0.13)	0.56 (0.40,0.69)

Abbreviation: area under receiver operating characteristic curve (AUC), net-reclassification of event (NRI$_e$), net-reclassification of non-event (NRI$_{ne}$)

Figure 1 10-year calibration plots in testing data sets

(A) Calibration plot of the MWCCS model
in the MWCCS validation data set

(B) Calibration plot of the MWCCS model
in the MACS validation data set

(C) Calibration plot of the MWCCS model
in the WIHS validation data set

(D) Calibration plot of the sex-specific models
in the MWCCS validation data set

(E) Calibration plot of the MACS sex-specific model
in the MACS validation data set

(F) Calibration plot of the WIHS sex-specific model
in the WIHS validation data set

(G) Calibration plot of the PCE
in the MWCCS validation data set

(H) Calibration plot of the PCE
in the MACS validation data set

(I) Calibration plot of the PCE
in the WIHS validation data set

5.3 Result

Table 2 presents the newly derived models. We also provided additional information on the tuning parameters from the boosting process and residual plots for evaluation of the sub-distributional proportional hazard assumption in the Appendix (Appendix C supplementary table 1 and supplementary figure 1 to 5, respectively). During the boosting steps, all predictors in the 5-year MWCCS model had coefficients of zero. Moreover, evaluation of for sub-distributional proportional hazard assumption suggested potential violation from the 5-year MACS model, the 5-year WIHS model, and the 10-year MACS model (Appendix C supplementary figures 1 to 3). Therefore, we only included details of the 10-year WIHS models and 10-year MWCCS model in table 2 and subsequent analysis.

Table 3 shows the details of testing data sets. Over 5 years, the percentage of non-HIV deaths to total deaths were 21.43%, 37.5%, and 30% for MACS, WIHS, and MWCCS testing data sets, respectively. Over 10 years, the corresponding percentage were 28.57%, 53.06%, and 45.71%. The average follow-up time suggested that most of the participants were followed to the end of each duration.

Table 4 illustrates predictive performance of the new models and the PCE. In the MACS testing data set, the MWCCS model had higher point estimation of AUC than that of PCE. Conversely, this measure from PCE was higher than those from all new models among the WIHS testing data set. From the MWCCS testing data set, both point estimations of AUC were roughly the same. Nevertheless, their 95% CI in each testing data set overlapped. Corresponding ROC curves are available in the Appendix (Appendix C supplementary figure 6).

Figure 1 illustrates calibration plots in testing data sets. These plots were evaluated based on the degree of alignment between the agreement line, which reflected the concordance between the predicted probability on the x-axis and the observed frequencies of events over the follow-up period on the y-axis, and the diagonal line. In the MACS testing data set, we could observe slightly better calibration from the MWCCS model (figure 1 plot B) as compared to the PCE (figure 1 plot E). In the WIHS testing data set, all plots seemed to have comparable calibration (figure 1 plots A, C, and F). In the MWCCS testing data set, MWCCS model (figure 1 plot D) appeared to have better calibration than the PCE (figure 1 plot G).

Table 4 also includes NRI_e and NRI_{ne} comparing each new model to the PCE. For the sex-specific model, we had positive NRI_e and negative NRI_{ne} in the WIHS testing data set. Thus, among WIHS participants, the risk prediction from the new models , as compared to the risk from the PCE, had positive net proportion of events assigned to the higher risk category and negative net proportion of non-events assigned to the lower risk category. For the MWCCS model, we obtained negative NRI_e and positive NRI_{ne} in the MACS testing data set. This finding suggests that the MWCCS model had negative net proportion of events assigned to the higher risk category and positive net proportion of non-events assigned to the lower risk category. In the WIHS testing data sets and the MWCCS testing data sets, both NRI_e and NRI_{ne} were positive. These patterns indicated that the new model assigned participants in either direction of risk group in a different degree.

5.4 Discussion

Our analysis supported that the new models had comparable predictive performance to the PCE for the prediction of non-HIV mortality over 10-year duration. All AUC had overlapping 95%CI. The majority of the agreement lines of calibration plots appeared to have similar length occupying the diagonal line. The NRI suggested some directionality of reclassification; however, the magnitude of this measures was small especially in the 10-year MWCCS testing data set.

The newly derived models exhibited some interesting features. First, we could see that the algorithmic-based model building does not always result in models that respect the required assumptions. In our case, some residual plots revealed the change in residual pattern over time with high magnitude; thus, we could not pursue further model evaluation from those models. Second, among the qualified models, age, sex, and race were included reflecting the importance of these demographic variable in the prediction. Third, PLHIV-specific blood tests ;which are CD4, CD8, and viral load; were selected from the algorithm. This finding supported our believe that PLHIV-specific variables should be chosen into the models. Fourth, some variables tended to have much larger size of coefficients than others such as CD4 and the interaction term between female sex (SEX0) and square root of viral load (Sqrt(VL)) in the 10-year MWCCS model. We might be able to infer that these variables play more an important role in this prediction.

Our study had a few limitations. First, the limited number of events hindered our ability to explore the models and their performance especially in subgroup analysis. Not only that the selection of only non-HIV death affected the available number of event but also the stratification by sex and race imposed this difficulty as well. Despite our effort in using long-standing

multicenter cohort studies, the proportion of non-HIV death appeared to be fewer than from the national surveillance. Compared to the US national surveillance data, the MWCCS has lower proportion of non-HIV deaths. Over five-year period, non-HIV death from the national surveillance was around 42% compared to 21.43% to 37.5% in our cohort[5]. This proportion in 2017 even went high up to 62.8% at the national level[4]. Moreover, this discrepancy could point towards the area that should further be included in the data collection site of the combined cohort studies to be more representative of the national level statistics. Second, some variables, such as IL6, were not included in our modeling process due to high proportions of missing values.

Despite these limitations, our study possessed several strengths. First, the MWCCS is a combined cohort study with long-term follow up which is an appropriate study design for building derivation of prediction model. Second, the utilized statistical model and algorithmic approach are more appropriate for the competing risk setting. Third, we offered a range of performance evaluation including quantitative evaluation by the AUC and the net reclassification index and qualitative evaluation which is the calibration plot.

To sum up, we set out to develop and evaluate a prediction model for non-HIV mortality from cardiovascular and PLHIV-specific risk factors with comparison to benchmarking models from the routine clinical practice. The discrimination performance of the new models was not significantly improved as compared to the PCE over 10-year period; however, we could observe some better calibration from the new models.

5.5 Public Health Implication

The newly derived models illustrated a slightly better calibration and some magnitude of risk re-classification. However, none of the models had sufficient discrimination. Therefore, we have highlighted the need to develop a model with higher discrimination to predict non-HIV mortality.

5.6 Conclusion

Our endeavor to build a prediction model for non-HIV mortality among PLHIV yielded a model with slightly improved calibration and some degree of re-classification as compared to the PCE. However, none of the new models demonstrated adequate predictive performance in the testing data sets.

5.7 Reference

1   Rank F, Alella JP, Nne A, Oorman CM, Oveless AOL, Ack J, *et al.* **Declining Morbidity**
        and Mortality among Patients with Advanced Human Immunodeficiency Virus
        Infection. *N Engl J Med* 1998; **338**:853.

2   Feinstein MJ, Bahiru E, Achenbach C, Longenecker CT, Hsue P, So-Armah K, *et al.*
        Patterns of cardiovascular mortality for HIV-infected adults in the United States:
        1999 to 2013. *Am J Cardiol* 2016; **117**:214–220.

3   Palella FJ, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, *et al.* **Mortality in**
        the Highly Active Antiretroviral Therapy Era. *JAIDS J Acquir Immune Defic Syndr*
        2006; **43**:27–34.

4   Bosh KA, Johnson AS, Hernandez AL, Prejean J, Taylor J, Wingard R, *et al.* **Vital Signs:**
        Deaths Among Persons with Diagnosed HIV Infection, United States, 2010–2018.
        MMWR 2020; **69**:1717–1724.

5   Adih WK, Selik RM, Hall HI, Surendera Babu A, Song R. **Associations and Trends in**
        Cause-Specific Rates of Death Among Persons Reported with HIV Infection, 23 U.S.
        Jurisdictions, Through 2011. *Open AIDS J* 2016; **10**:144–145.

6   Tate JP, Justice AC, Hughes MD, Bonnet F, Reiss P, Mocroft A, *et al.* **An internationally**
        generalizable risk index for mortality after one year of antiretroviral therapy. *AIDS*
        2013; **27**:563–572.

7   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. *J Acquir Immune Defic Syndr* 2013; **62**:149–163.

8   Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, *et al.* **2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the American college of cardiology/American heart association task force on practice guidelines**. *J Am Coll Cardiol* 2014; **63**:2889–2934.

9   Fine JP, Gray RJ. **A Proportional Hazards Model for the Subdistribution of a Competing Risk**. *J Am Stat Assoc* 1999; **94**:496–509.

10   Austin PC, Lee DS, Fine JP. **Introduction to the Analysis of Survival Data in the Presence of Competing Risks**. *Circulation* 2016; **133**:601–609.

11   Austin PC, Fine JP. **Practical recommendations for reporting Fine-Gray model analyses for competing risk data**. Published Online First: 2017. doi:10.1002/sim.7501

12   Bi Q, Goodman KE, Kaminsky J, Lessler J. **Practice of Epidemiology What is Machine Learning? A Primer for the Epidemiologist**. *Am J Epidemiol* 2019; **188**:222–2239.

13   Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR. **The Multicenter AIDS Cohort Study: Rationale, Organization, and Selected Characteristics of the Participant**. *Am J Epidemiol* 1987; **126**:310–318.

14   Detels R, Jacobson L, Margolick J, Martinez-Maza O, Muñoz A, Phair J, *et al.* **The multicenter AIDS Cohort Study, 1983 to ...** *Public Health* 2012; **126**:196–198.

15   Barkan SE, Melnick SL, Preston-martin S, Weber K, Kalish LA, Miotti P, *et al.* **The**

Women' s Interagency HIV Study. *Epidemiology* 1998; **9**:117–125.

16   Bacon MC, Von Wyl V, Alden C, Sharp G, Robison E, Hessol N, *et al.* **The Women's Interagency HIV Study: an observational cohort brings clinical sciences to the bench**. *Clin Diagn Lab Immunol* 2005; **12**:1013–1019.

17   Adimora AA, Ramirez C, Benning L, Greenblatt RM, Kempf M-C, Tien PC, *et al.* **Cohort Profile: The Women's Interagency HIV Study (WIHS)**. *Int J Epidemiol* 2018; **47**:393-394i.

18   Wong ND, Amsterdam EA, Blumenthal RS. *ASPC Manual of Preventive Cardiology*. New York: Demos Medical Publishing, LLC; 2015.

19   Patel AA, Budoff MJ. **Coronary Artery Disease in Patients with HIV Infection: An Update**. *Am J Cardiovasc Drugs* 2021; **21**:411–417.

20   R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

21   Harald Binder (2021). CoxBoost: Cox Models by Likelihood Based Boosting for a Single Survival Endpoint or Competing Risks. R package version 1.5.

22   Li J, Thomas ·, Scheike H, Zhang M-J, Li J, Zhang · M.-J, *et al.* **Checking Fine and Gray subdistribution hazards model with cumulative sums of residuals**. *Lifetime Data Anal* 2015; **21**:197–217.

23   Paul Blanche, Jean-Francois Dartigues, Helene Jacqmin-Gadda (2013). **Estimating and Comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks**. Statistics in Medicine, 32(30), 5381-5397.

URL http://onlinelibrary.wiley.com/doi/10.1002/sim.5958/full .

24   Ulla B. Mogensen, Hemant Ishwaran, Thomas A. Gerds (2012). **Evaluating Random Forests for Survival Analysis Using Prediction Error Curves**. Journal of Statistical Software, 50(11), 1-23. URL https://www.jstatsoft.org/v50/i11.

25   Eisuke Inoue (2018). nricens: NRI for Risk Prediction Models with Time to Event and Binary Response Data. R package version 1.6. https://CRAN.R-project.org/package=nricens

CHAPTER 6

Summary

6.1 Summary

This dissertation presented the findings on the development of prediction models for new subclinical atherosclerotic plaque, all-cause mortality, and non-HIV mortality among people living with HIV (PLHIV) based on cardiovascular risk factors. We utilized the MWCCS, a combined study from two long-standing cohort studies for PLHIV in the US, to generate models for the overall population and sex-specific models for each outcome. We also incorporated modern statistical techniques, including penalization and ensemble, to maximize predictive accuracy and avoid overfitting.

From Chapters 3 to 5, the new models appear to have better performance than benchmarking models, at least for one of the performance assessments. The net reclassification also supports the more appropriate direction of risk group classification using the new model. However, only the 10-year sex-specific and MWCCS model for all-cause mortality had superior discrimination compared to the PCE. The superior performance in the new models results from the model building in the PLHIV group with more PLHIV-specific covariates, using both continuous and categorical formats for each predictor interaction terms, and implementation of more advance statistical methods.

6.2 Limitations and challenge

This dissertation has a number of challenges. Most importantly, model building requires a large number of events that might not be sufficient if we stratify our data set too thin. Thus, we only derive the model from the whole data set and stratified by sex. We could have evaluated the sex-race-specific model to refine predictions with more events. Second, some parts of model

building, and evaluation were subjective or based on arbitrary cut-off value, including the nonlinearity test, the calibration plots, and the net reclassification index. Third, the modeling methods are flexible, which might need further exploration. For example, the variable selection during the elastic net regularization could have been done in a group-selection fashion rather than to treat each predictor independently. This different choice of modeling process will come with further justification on the model.

Despite these limitations and challenges, this work has illustrated model building methods that result in models with better performance and a model that has a potential for external validation. Due to the continuous increase of non-communicable diseases among people living with HIV, we would encourage further development in a predictive model to accommodate clinical decision-making in this era.

APPENDIX A

Supplementary materials to Chapter 3

Supplementary Table 1 Tuning parameters of newly derived models from the training data sets

MACS sex-specific model		WIHS sex-specific model		MWCCS model	
Alpha	Lambda 1 SE	Alpha	Lambda 1 SE	Alpha	Lambda 1 SE
0.729	0.0882	0.001	44.28	0.0252	1.00

APPENDIX B

Supplementary materials to Chapter 4

Supplementary Table 1 Tuning parameters of newly derived models from the training data sets

Duration	MACS sex-specific model		WIHS sex-specific model		MWCCS model	
	Alpha	Lambda 1 SE	Alpha	Lambda 1 SE	Alpha	Lambda 1SE
5-year duration	1.0000	0.0117	1.0000	0.0221	0.7290	0.0230
10-year duration	1.0000	0.0087	1.000	0.0304	1.0000	0.0198

Supplementary figure 1 Scaled Schoenfeld residuals plot

for the 5-year MACS sex-specific model
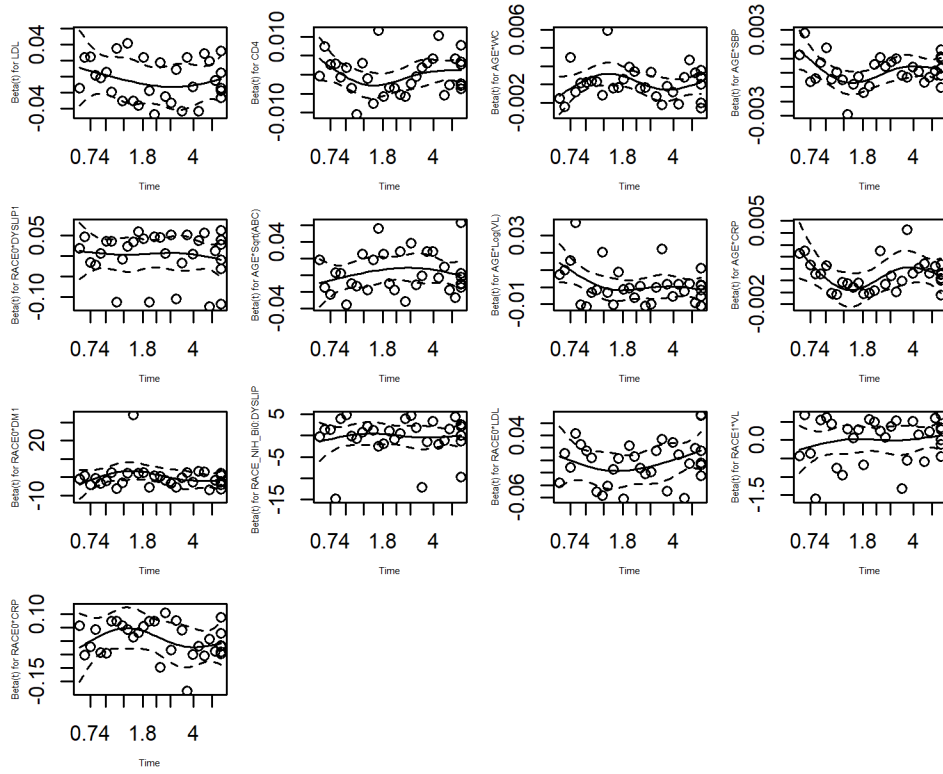
Supplementary figure 2 Scaled Schoenfeld residuals plot

for the 5-year WIHS sex-specific model
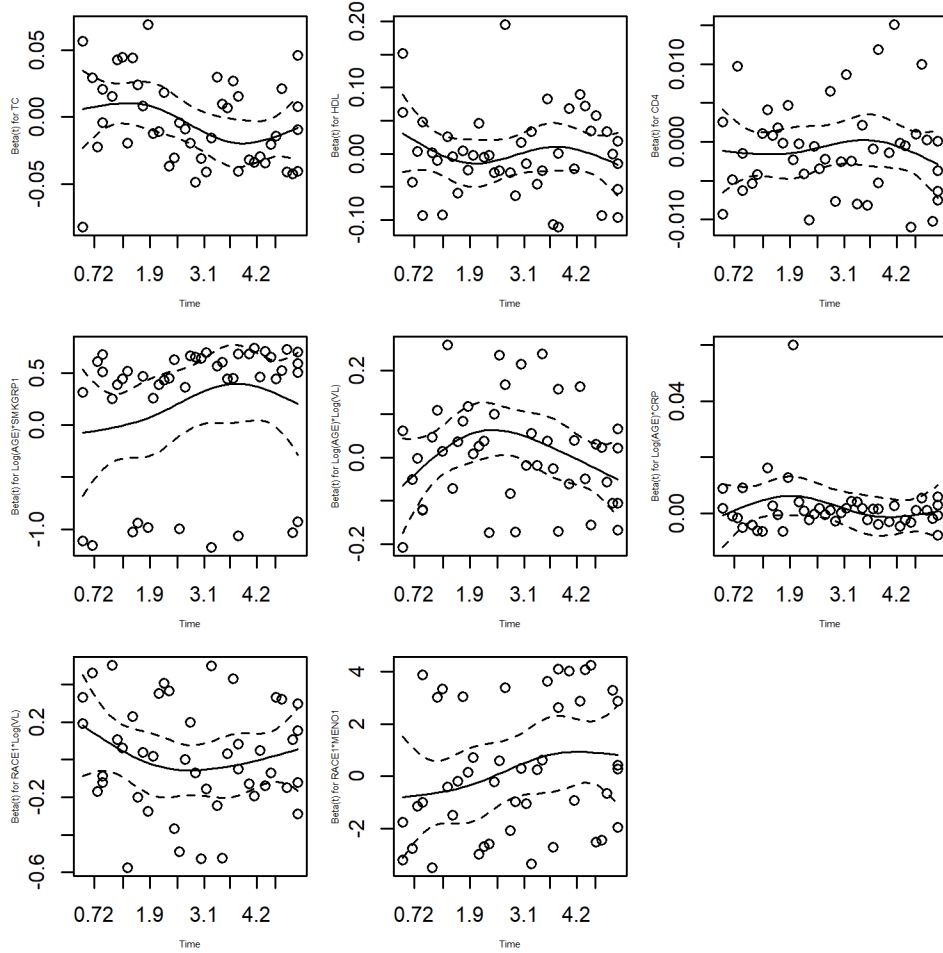
Supplementary figure 3 Scaled Schoenfeld residuals plot for the 5-year MWCCS model
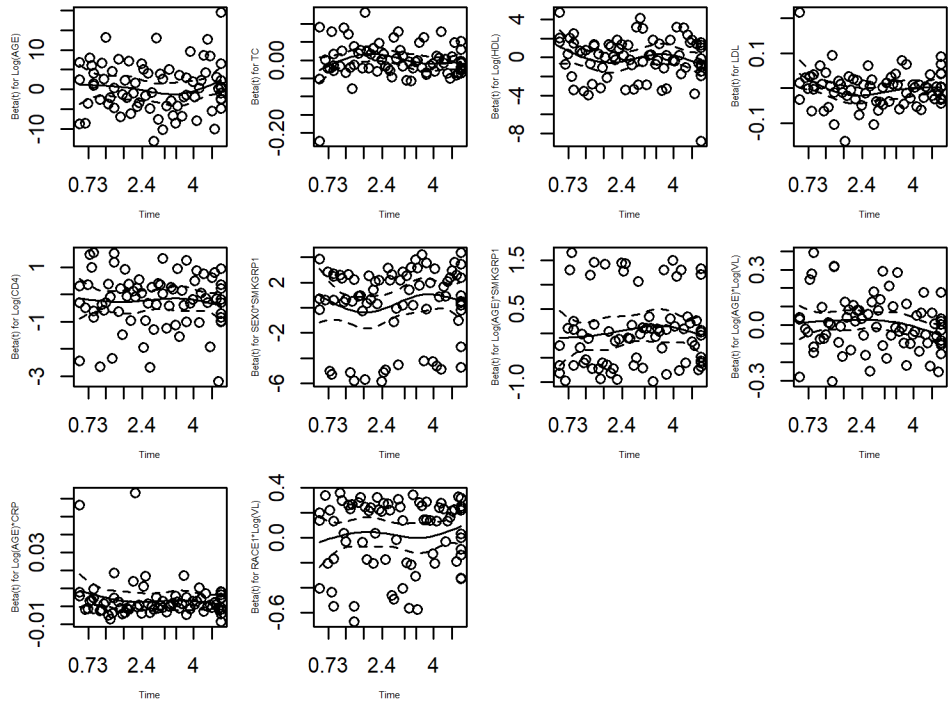
Supplementary figure 4 Scaled Schoenfeld residuals plot

for the10-year MACS sex-specific model
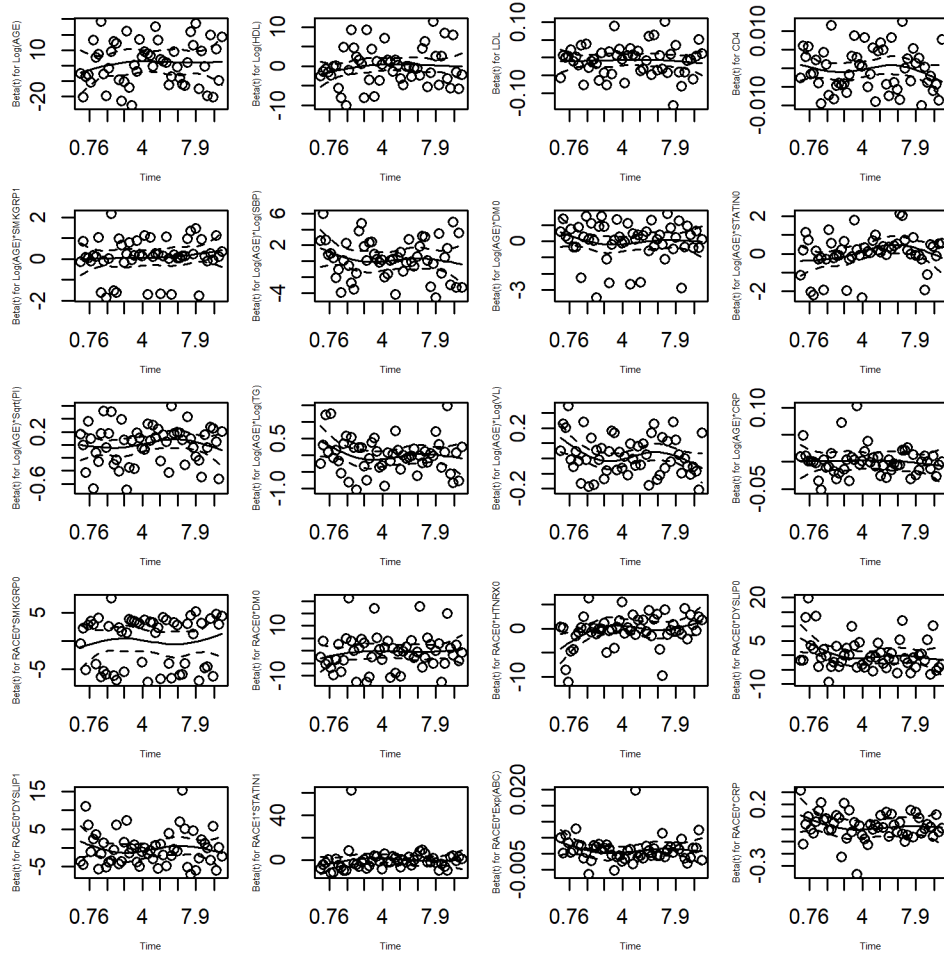
Supplementary figure 5 Scaled Schoenfeld residuals plot

for the10-year WIHS sex-specific model
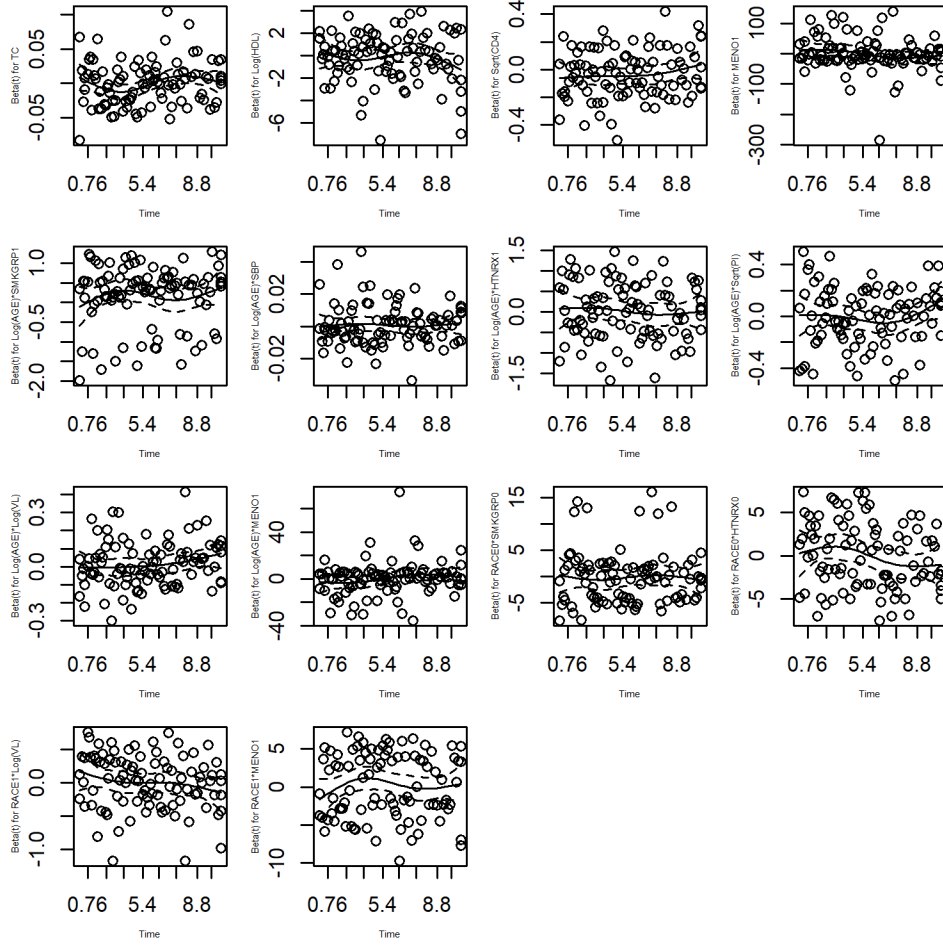
Supplementary figure 6 Scaled Schoenfeld residuals plot

for the10-year MWCCS model
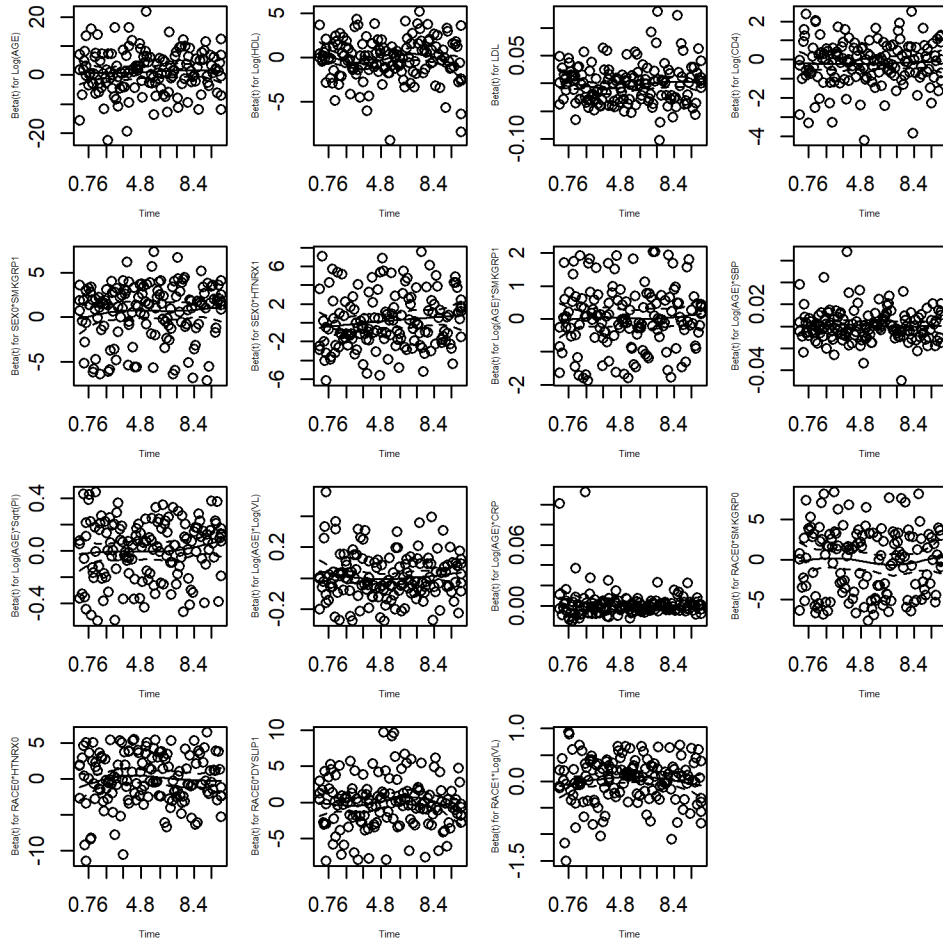
Supplementary text: example of an individual risk calculation based on the 10-year MWCCS model

We will apply the model to calculate 10-year mortality risk for a 55-year white PLHIV who is actively smoking. His SBP is 138 mmHg with no antihypertensive treatment. He has dyslipidemia with HDL of 74.5 mg/dL and LDL of 111 mg/dL. His HIV blood test reveals a CD4 level of 882 cell/mm^3 and a viral load of 50 copies/ ml. He has never taken PI medication. His CRP blood level is 0.7 mg/dL.

Step 1: Multiplication of the individual predictor values and the regression coefficients

We can write the product "A" described in the footnote of table 2 as follows.

$$
\begin{aligned}
A = &\left((1.1507) * Log(AGE)\right) + \left((-1.394 * 10^{-1}) * Log(HDL)\right) + \left((-1.7487 * 10^{-3}) * LDL\right) \\
&+ \left((-1.9059 * 10^{-1}) * Log(CD4)\right) + \left((6.5809 * 10^{-1}) * SEX0 * SMKGRP1\right) \\
&+ \left((1.7165 * 10^{-1}) * SEX0 * HTNRX1\right) \\
&+ \left((3.6563 * 10^{-2}) * Log(AGE) * SMKGRP1\right) \\
&+ \left((6.6744 * 10^{-4}) * Log(AGE) * SBP\right) \\
&+ \left((4.8305 * 10^{-3}) * Log(AGE) * Sqrt(PI)\right) \\
&+ \left((1.2417 * 10^{-2}) * Log(AGE) * Log(VL)\right) \\
&+ \left((1.5154 * 10^{-3}) * Log(AGE) * CRP\right) \\
&+ \left((-1.5278 * 10^{-1}) * RACE0 * SMKGRP0\right) \\
&+ \left((-3.3292 * 10^{-2}) * RACE0 * HTNRX0\right) \\
&+ \left((-7.1972 * 10^{-2}) * RACE0 * DYSLIP1\right) \\
&+ ((2.0527 * 10^{-2}) * RACE1 * Log(VL))
\end{aligned}
$$

Then, we replace the predictor values with those from the given individual profile.

$$A = \big((1.1507) * Log(55)\big) + \big((-1.394 * 10^{-1}) * Log(74.5)\big) + \big((-1.7487 * 10^{-3}) * 111\big)$$

$$+ \big((-1.9059 * 10^{-1}) * Log(882)\big) + \big((6.5809 * 10^{-1}) * 0 * 1\big)$$

$$+ \big((1.7165 * 10^{-1}) * 0 * 0\big) + \big((3.6563 * 10^{-2}) * Log(55) * 1\big)$$

$$+ \big((6.6744 * 10^{-4}) * Log(55) * 138\big)$$

$$+ \big((4.8305 * 10^{-3}) * Log(55) * Sqrt(0)\big)$$

$$+ \big((1.2417 * 10^{-2}) * Log(55) * Log(50)\big)$$

$$+ \big((1.5154 * 10^{-3}) * Log(55) * 0.7\big) + \big((-1.5278 * 10^{-1}) * 1 * 0\big)$$

$$+ \big((-3.3292 * 10^{-2}) * 1 * 1\big) + \big((-7.1972 * 10^{-2}) * 1 * 1\big)$$

$$+ \big((2.0527 * 10^{-2}) * 0 * Log(50)\big)$$

$A = 2.9864$

Step 2 Compute B from the exponentiation of A
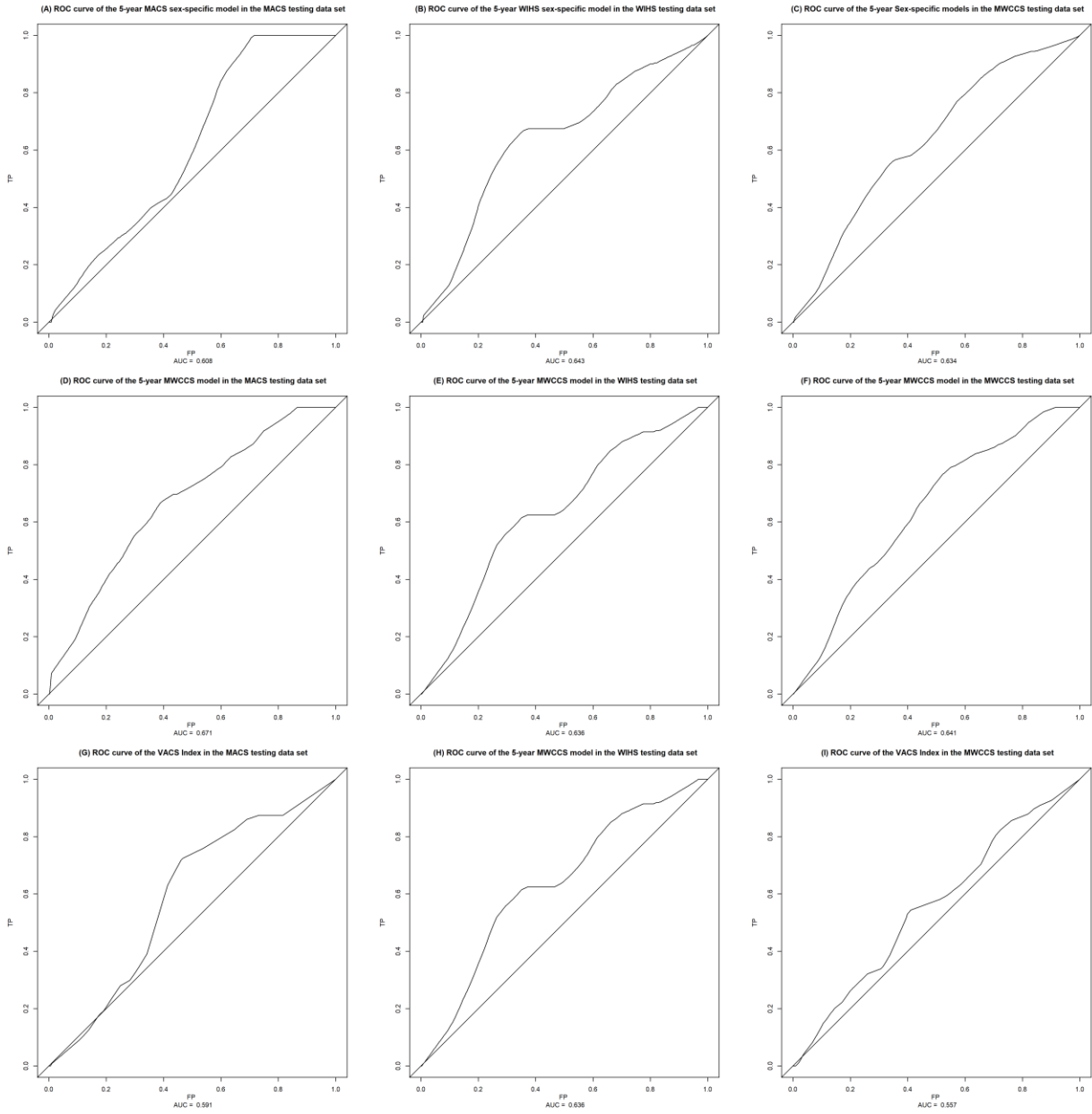
$B = Exp(A) = Exp(2.9864) = 19.8142$

Step 3 Calculate the 10-year risk of death

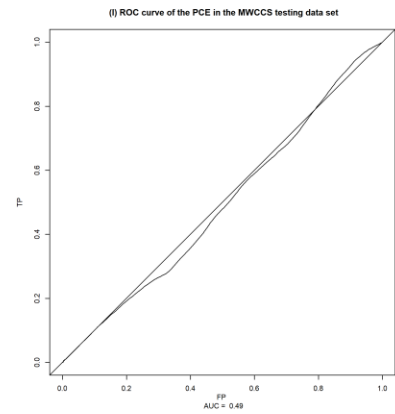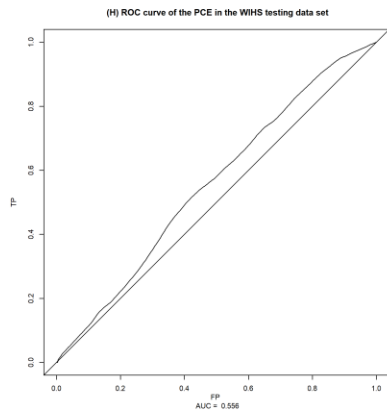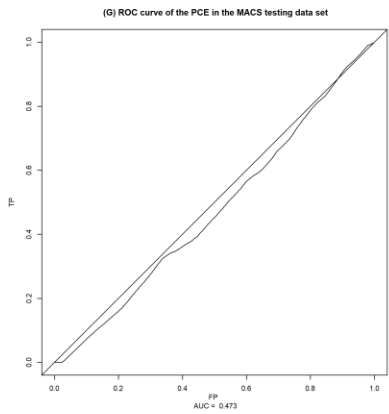From Table 2, the baseline survival probability is 0.9953. We compute the risk of death over 10-year as follows.
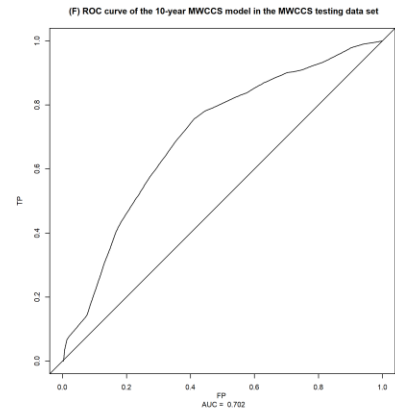
$$10 - year\ risk = 1 - S_0(10)^B = 1 - 0.9953^{19.8142} = 1 - 0.9109 = 0.0891 \approx 9\%$$

Over a 10-year period, the risk of death due to any cause of this individual is approximately 9% based on the 10-year MWCCS model.

Supplementary figure 7 5-year ROC curves in testing data sets
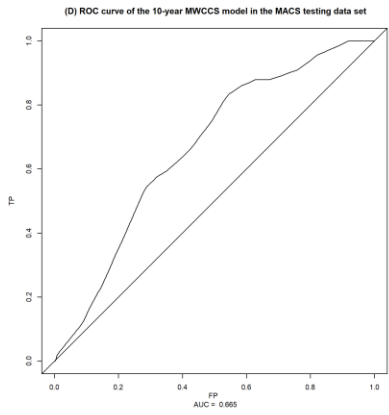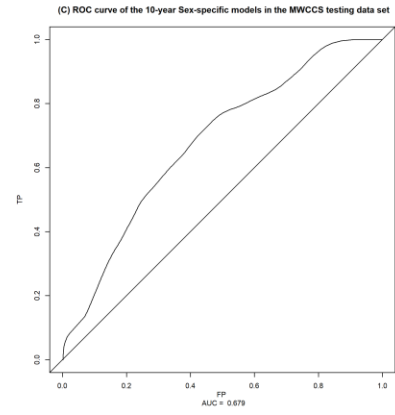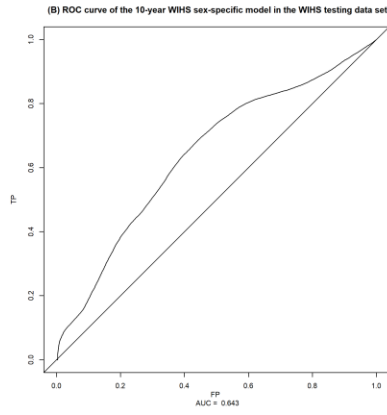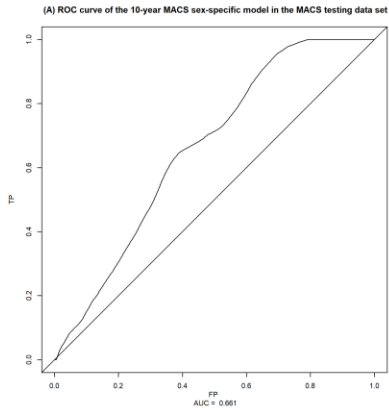

(A) ROC curve of the 5-year MACS sex-specific model in the MACS testing data set
AUC = 0.608

(B) ROC curve of the 5-year WIHS sex-specific model in the WIHS testing data set
AUC = 0.643

(C) ROC curve of the 5-year Sex-specific models in the MWCCS testing data set
AUC = 0.634

(D) ROC curve of the 5-year MWCCS model in the MACS testing data set
AUC = 0.671

(E) ROC curve of the 5-year MWCCS model in the WIHS testing data set
AUC = 0.636

(F) ROC curve of the 5-year MWCCS model in the MWCCS testing data set
AUC = 0.641

(G) ROC curve of the VACS Index in the MACS testing data set
AUC = 0.591

(H) ROC curve of the 5-year MWCCS model in the WIHS testing data set
AUC = 0.636

(I) ROC curve of the VACS Index in the MWCCS testing data set
AUC = 0.557

Supplementary figure 8 10-year ROC curves in testing data sets

(A) ROC curve of the 10-year MACS sex-specific model in the MACS testing data set
AUC = 0.661

(B) ROC curve of the 10-year WIHS sex-specific model in the WIHS testing data set
AUC = 0.643

(C) ROC curve of the 10-year Sex-specific models in the MWCCS testing data set
AUC = 0.679

(D) ROC curve of the 10-year MWCCS model in the MACS testing data set
AUC = 0.665

(E) ROC curve of the 10-year MWCCS model in the WIHS testing data set
AUC = 0.677

(F) ROC curve of the 10-year MWCCS model in the MWCCS testing data set
AUC = 0.702

(G) ROC curve of the PCE in the MACS testing data set
AUC = 0.473

(H) ROC curve of the PCE in the WIHS testing data set
AUC = 0.556

(I) ROC curve of the PCE in the MWCCS testing data set
AUC = 0.49

APPENDIX C

**Supplementary materials to Chapter 5**

Supplementary Table 1 Tuning parameters of newly derived models from the training data sets

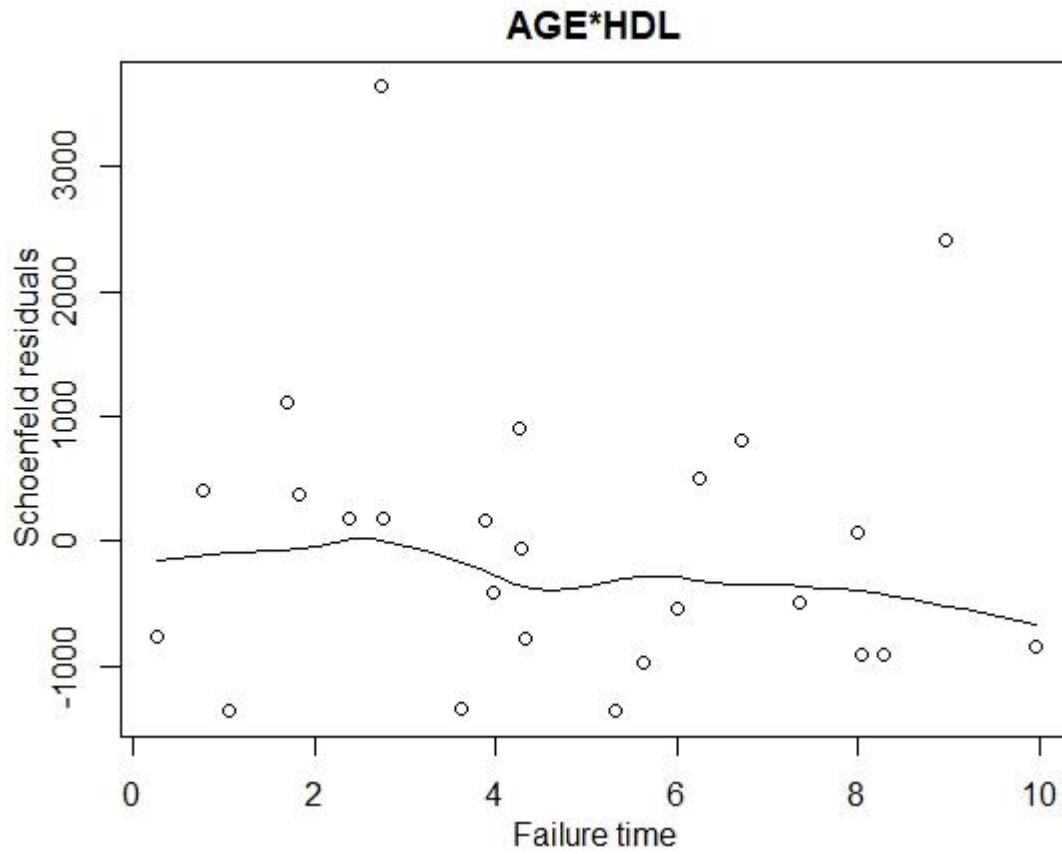| Duration	MACS sex-specific model		WIHS sex-specific model		MWCCS model	
	Boost steps	Partial log-likelihood	Boost steps	Partial log-likelihood	Boost steps	Partial log-likelihood
5-year duration	34	-72.2315	18	-96.7470	NA	NA
10-year duration	3	-160.4416	72	-248.0718	42	-459.5592

Supplementary Figure 1 Scaled Schoenfeld residual plots

for the 5-year MACS sex-specific model
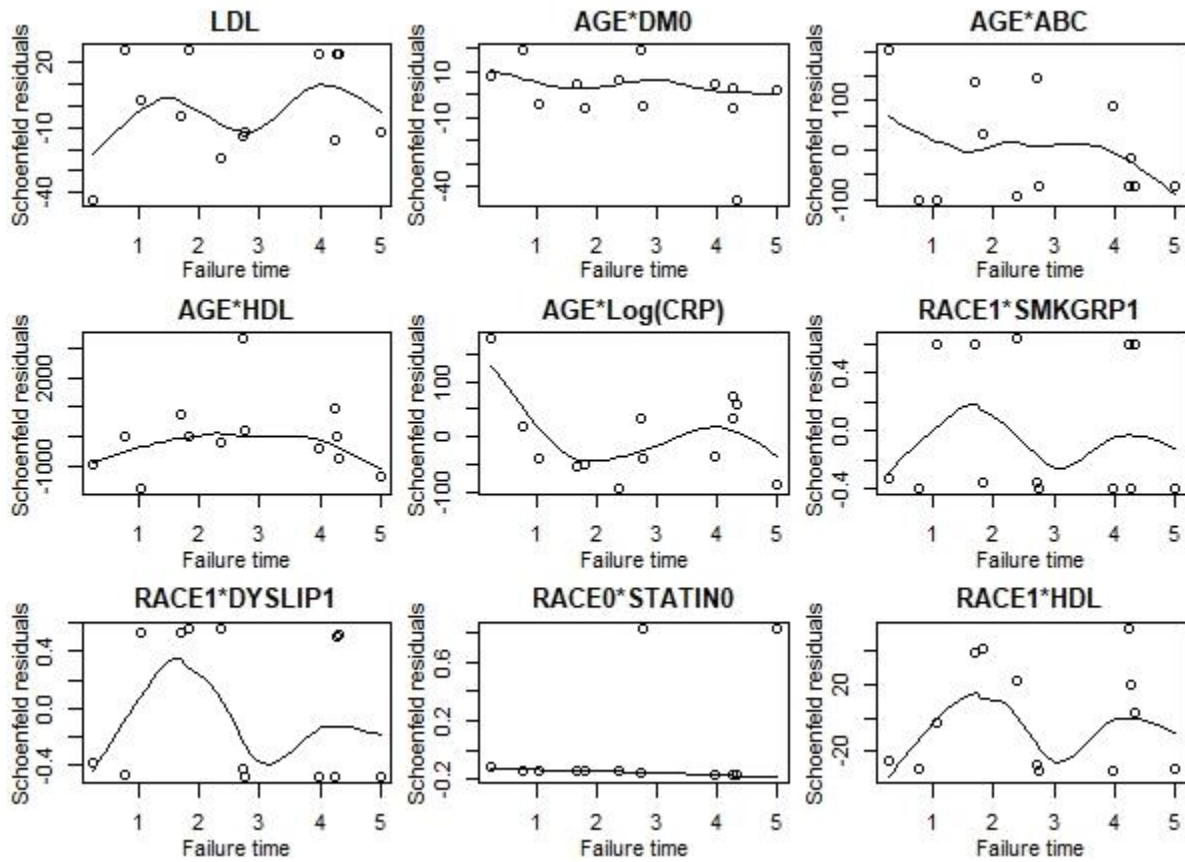
Supplementary Figure 2 Scaled Schoenfeld residual plots

for the 5-year WIHS sex-specific model
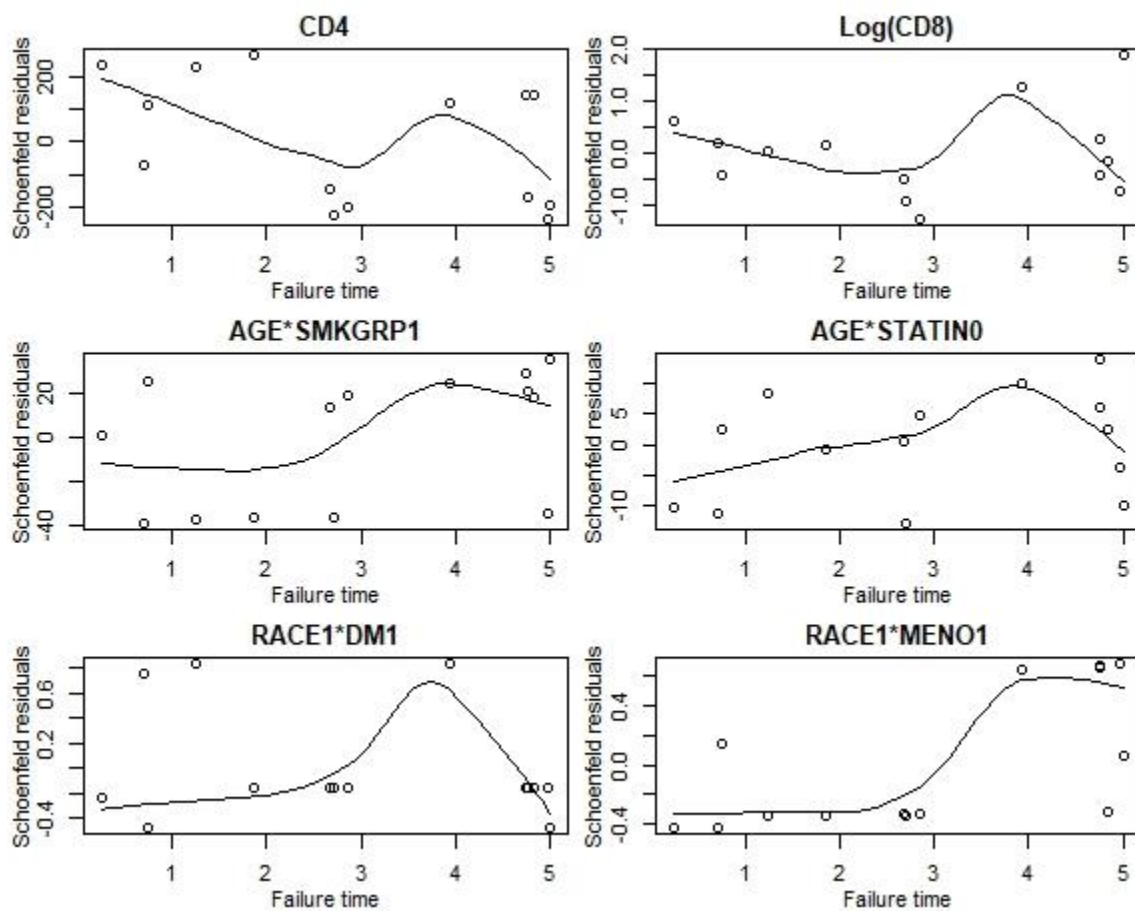
Supplementary Figure 3 Scaled Schoenfeld residual plots

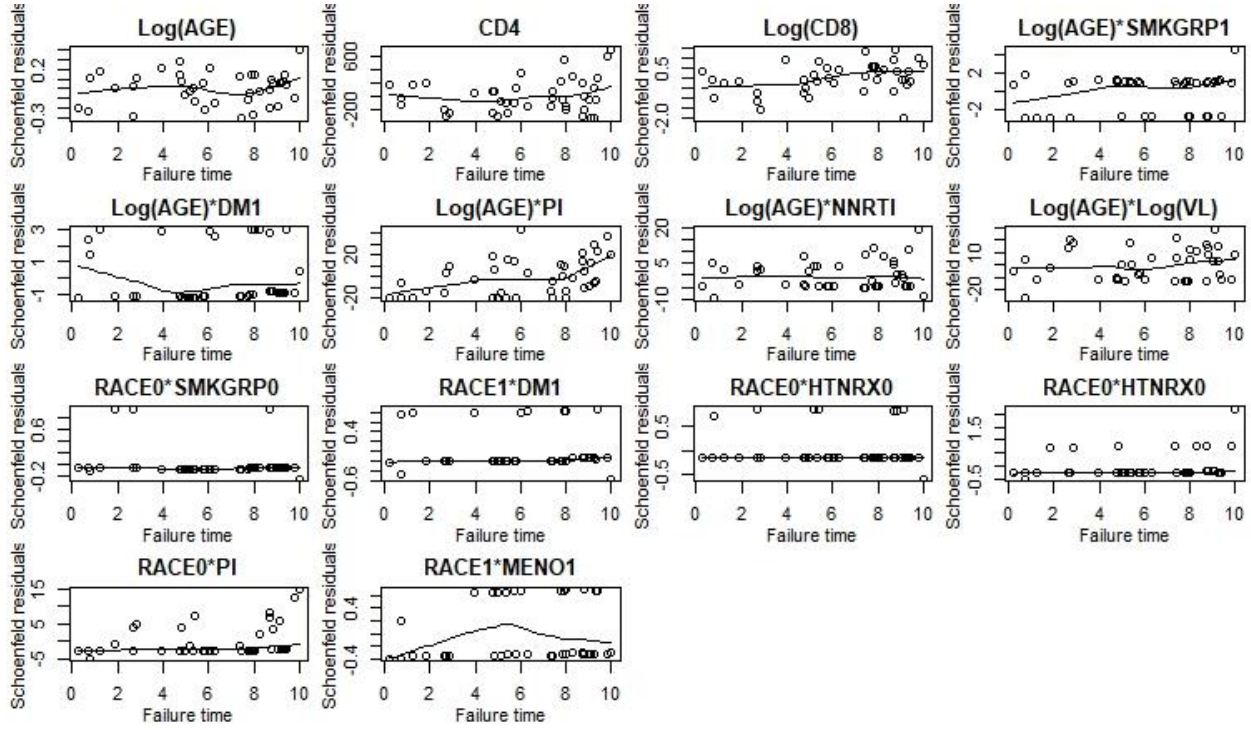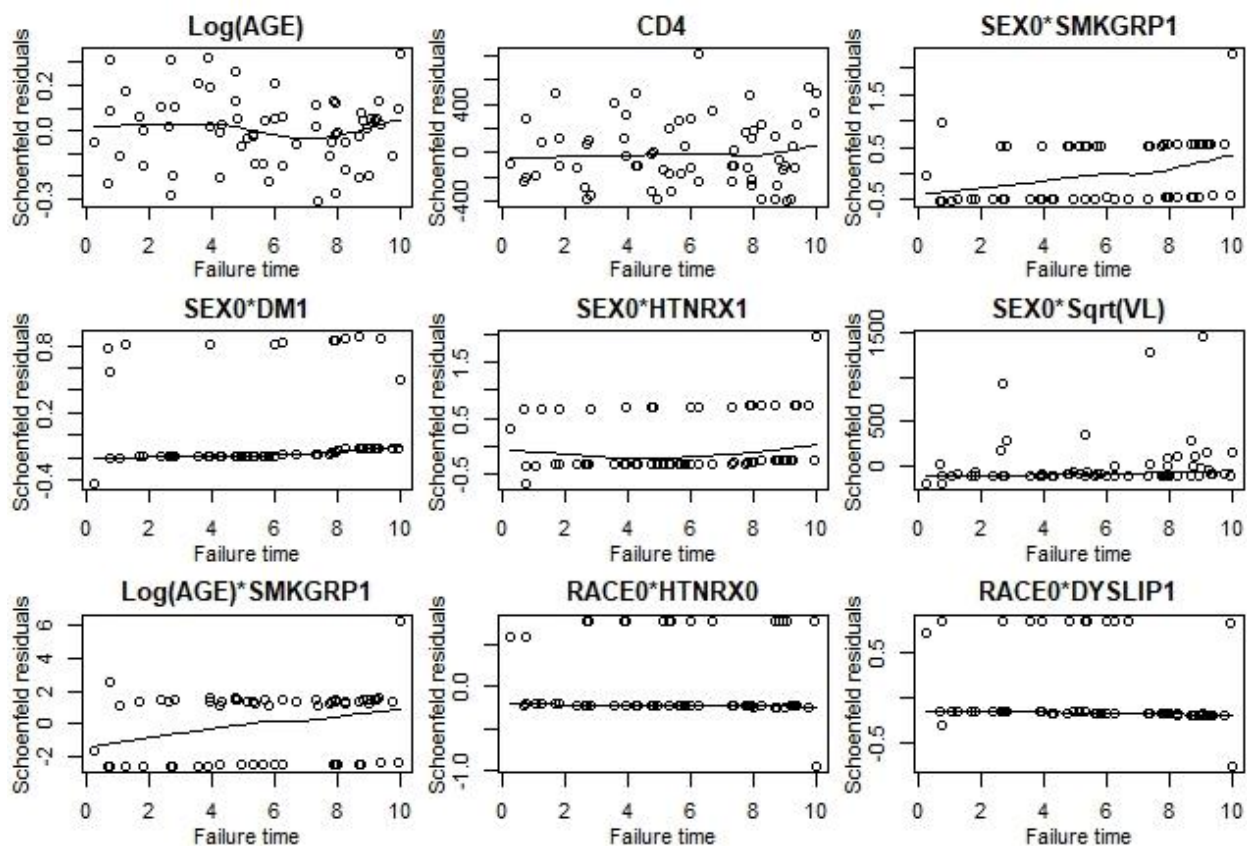**for the 10-year MACS sex-specific model**

AGE*HDL

Supplementary Figure 4 Scaled Schoenfeld residual plots

for the 10-year WIHS sex-specific model

Supplementary Figure 5 Scaled Schoenfeld residual plots

for the 10-year MWCCS sex-specific model

Supplementary figure 6 10-year ROC curves in testing data sets

(A) ROC curve of the WIHS sex-specific model in the WIHS testing data set

(B) ROC curve of the MWCCS model in the MACS testing data set

(C) ROC curve of the MWCCS model in the WIHS testing data set

(D) ROC curve of the MWCCS model in the MWCCS testing data set

(E) ROC curve of the PCE in the MACS testing data set

(F) ROC curve of the PCE in the WIHS testing data set

(G) ROC curve of the PCE in the MWCCS testing data set