The Influence of Selection Bias on Racial Differences in Reproductive Aging and Accelerated Health Declines in the Study of Women's Health Across the Nation

by

Alexis Reeves

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Epidemiological Science)
in the University of Michigan
2021

Doctoral Committee:

Professor Siobán D. Harlow, Chair
Professor Michael R. Elliott
Professor William H. Herman
Assistant Professor Carrie A. Karvonen-Gutierrez
Associate Professor Tené E. Lewis, Emory University

Alexis Reeves

alexisnr@umich.edu

ORCID iD: 0000-0002-6831-7173

Dedication

Dedicated in loving memory to Dorothea “Nan” Reeves.

The giant’s shoulders from which all of us Reeves reach for the stars.

June 2, 1927 – October 14, 2020

Acknowledgements

I feel honored and privileged to get to do this work every day. There are so many people that were in my corner from early in my career to now, making this dissertation possible – thank you to all seen and unseen that helped me walk this path.

First, thank you to the Study of Women's Health Across the Nation's (SWAN) dedicated investigators, study staff, and chiefly, the participants who have made this wonderful and rich cohort possible. None of this work would be realized without your dedication.

Thank you to Dr. Siobán Harlow, my committee chair, advisor, and lifelong mentor. I am so grateful and honored to be your last doctoral student. You have inspired and encouraged me to be strong in my values and have honed my voice as a researcher. This dissertation would surely be impossible without your unwavering support and mentorship.

Thank you to my committee members: Dr. Michael Elliott for your mentorship and brilliant contributions to the analysis used in this dissertation. Dr. William Herman for opening doors to new opportunities and always providing encouragement through hard times. Dr. Carrie Karvonen-Gutierrez for your mentorship, thoughtfulness, and irreplaceable knowledge of the women and history of the SWAN cohort. And Dr. Tené Lewis, for your constant encouragement, mentorship, and being a part of this work despite having to travel to meetings!

I would also like to extend my gratitude to our comps study group (Aliya, Steph and Traci) for unwavering support, happy hours, and laughs. To Viktoryia for our causal inference and running

chats. And to Dr. Lauren Brown, for the thoughtful and real advice, friendship and commiseration over shared trials and tribulations. To Dr. Kirsten Herold, for our lively chats and critique of my writing throughout this process. Also, to the faculty, staff, and students at the Center for Midlife Sciences for providing a research "home" for this work and me.

Thank you to my husband Kevin for your unwavering support throughout this process, for encouraging me when I was down, for supporting my mental health and well-being, and celebrating all the small and big wins along the way. And finally, to my parents, you have given me a foundation to thrive - thank you for your unconditional love and belief in me.

There are numerous funding sources to thank in supporting this dissertation. Thank you to Siobán Harlow, the Mary Fran Sowers Fund at the Center for Midlife Sciences, the UM Epidemiology Department, and Rackham Graduate School for supporting me in this work. This work was also supported by grants from the National Institutes of Health including:

The National Institute on Aging's Ruth L. Kirschstein National Research Service Award (NRSA) Individual Pre-doctoral Fellowship to Promote Diversity in Health-Related Research (1F31AG064856-01).

SWAN has grant support from the National Institutes of Health (NIH), DHHS, through the National Institute on Aging (NIA), the National Institute of Nursing Research (NINR) and the NIH Office of Research on Women's Health (ORWH) (Grants U01NR004061; U01AG012505, U01AG012535, U01AG012531, U01AG012539, U01AG012546, U01AG012553, U01AG012554, U01AG012495, and U19AG063720). The content of this dissertation is solely the responsibility of the author and does not necessarily represent the official views of the NIA, NINR, ORWH or the NIH.

Clinical Centers: *University of Michigan, Ann Arbor – Siobán Harlow, PI 2011 – present, MaryFran Sowers, PI 1994-2011*; *Massachusetts General Hospital, Boston, MA – Sherri-Ann Burnett-Bowie, PI 2020 – Present; Joel Finkelstein, PI 1999 – 2020*; *Robert Neer, PI 1994 – 1999; Rush University, Rush University Medical Center, Chicago, IL – Imke Janssen, PI 2020 – Present; Howard Kravitz, PI 2009 – 2020*; *Lynda Powell, PI 1994 – 2009; University of California, Davis/Kaiser – Elaine Waetjen and Monique Hedderson, PIs 2020 – Present; Ellen Gold, PI 1994 - 2020*; *University of California, Los Angeles – Arun Karlamangla, PI 2020 – Present; Gail Greendale, PI 1994 - 2020*; *Albert Einstein College of Medicine, Bronx, NY – Carol Derby, PI 2011 – present, Rachel Wildman, PI 2010 – 2011; Nanette Santoro, PI 2004 – 2010; University of Medicine and Dentistry – New Jersey Medical School, Newark – Gerson Weiss, PI 1994 – 2004;* and the *University of Pittsburgh, Pittsburgh, PA – Rebecca Thurston, PI 2020 – Present; Karen Matthews, PI 1994 - 2020.*

NIH Program Office: *National Institute on Aging, Bethesda, MD – Rosaly Correa-de-Araujo 2020 - present; Chhanda Dutta 2016- present; Winifred Rossi 2012–2016; Sherry Sherman 1994 – 2012; Marcia Ory 1994 – 2001; National Institute of Nursing Research, Bethesda, MD – Program Officers.*

Central Laboratory: *University of Michigan, Ann Arbor – Daniel McConnell* (Central Ligand Assay Satellite Services).

Coordinating Center: *University of Pittsburgh, Pittsburgh, PA – Maria Mori Brooks, PI 2012 - present; Kim Sutton-Tyrrell, PI 2001 – 2012; New England Research Institutes, Watertown, MA - Sonja McKinlay, PI 1995 – 2001.*

Steering Committee: Susan Johnson, Current Chair
Chris Gallagher, Former Chair

Table of Contents

List of Tables

List of Figures

Abstract

Aging cohort studies recruit participants at an age before most of the population will experience the study outcome to document its natural history and related risk factors. Usually assumptions regarding "normative" aging, including average age at onset, are based on data on White populations. However, evidence suggests that Black and Hispanic populations experience "weathering" or accelerated health declines compared to Whites due to cumulative experience of social, economic, and political marginalization. When population subgroups have a differential probability of inclusion, selection bias, i.e., a distortion caused by differential selection into or out of a study, will likely misinform understanding of aging in these populations.

By accounting for left truncation and left and right censoring in the Study of Women's Health Across the Nation (SWAN), this dissertation addresses the main forms of selection bias in cohort studies and calculates their effects on estimates of racial/ethnic differences in reproductive aging and onset of cardio-metabolic risk. SWAN, a 25-year longitudinal cohort of midlife women, utilized a cross-sectional screener to recruit a cohort of Black, Chinese, Hispanic, Japanese, and White women.

Aim 1 created a framework for examining selection bias at study commencement and quantified how selection mechanisms from the cross-sectional screening study into the cohort biased estimates of racial/ethnic differences. Two main mechanisms were identified, *eligibility* and *participation*. Black and Hispanic women had the lowest eligibility rates stemming from high rates of surgical menopause. Their eligibility rates decreased with increasing age at a higher rate than in White women. Lower education was associated with lower eligibility for White women only. Participation was associated with demographic characteristics with no evidence of a "healthy volunteer" bias. Failure to account for selection at study commencement may underestimate racial/ethnic disparities in health, especially for Black and Hispanic women.

Aim 2 assessed the effects of selection bias on estimates of racial/ethnic differences in age at the final menstrual period (FMP). Results suggest that previous analyses that did not fully account for selection bias via left truncation and right censoring overestimated average age of FMP in Black and Hispanic women. In Black women, age of natural FMP was overestimated by an average of 1.1 years. Adjusting for selection bias, Black and Hispanic women had earlier natural and surgical FMPs compared to White women independent of socioeconomic factors and health behaviors.

Aim 3 assessed the effects of selection bias on estimates of racial/ethnic differences in the timing of cardio-metabolic outcomes including hypertension, isolated systolic hypertension, insulin resistance and diabetes. Selection at study commencement had greater effects on outcomes with earlier age at onset (hypertension), whereas selection via loss to follow up had greater effects on outcomes with later onsets (insulin resistance, diabetes). Addressing selection bias via left truncation, left and right censoring led to an average 20-year decrease in predicted median age of onset in cardio-metabolic outcomes but decreased the predicted difference in age at onset for Black and Hispanic women compared to White women. However, Black and Hispanic women still had significantly earlier onset of each outcome.

By providing a framework to account for and methods to mitigate selection bias, this dissertation documented accelerated health declines in Black and Hispanic women in SWAN across multiple reproductive and cardio-metabolic outcomes – highlighting the need for new and continued research/interventions aimed at the structural causes of racial disparities in health.

Chapter 1 Introduction

1.1 Overview

Cohorts of aging aim to understand the natural history of disease, assessing the causes and consequences of major drivers of disease, disability, and shortened lifespan[1–3]. To accomplish this goal, cohorts are assembled to observe the natural course of a life transition or onset of a disease by recruiting participants at a calendar age before most of the population has experienced the event of interest[2,3]. Historically, many of the hallmark cohorts of aging have been assembled for and by the hegemonic racial demographic[1] and thus the standard for normative aging has been in White populations. However, an emerging body of research suggests that Black and Hispanic women "weather" or experience accelerated health declines compared to White women in the US, potentially explaining persistent racial differences in lifespan[4–9]. As more cohorts aim to understand normative aging in multiple racial/ethnic groups[10–16] issues of selection bias into the cohort require more consideration. Most publications to date have relegated the issue of selection bias, especially at study commencement, to a short mention in the limitations section; although selection may particularly bias aging research for Black and Hispanic racial groups. This practice leaves a critical gap in scientific knowledge regarding how the methods used to select and analyze cohorts may be biased, potentially impeding an accurate understanding of aging in minoritized groups.

The overarching goal of this dissertation was to account for and examine the effect of three main forms of selection bias, left truncation, left censoring and right censoring – found in cohorts of aging on racial/ethnic differences in reproductive aging and onset of cardio-metabolic risk in the Study of Women's Health Across the Nation (SWAN). Given evidence of "weathering" or accelerated aging and health declines among minoritized populations[4–9], the main hypothesis was that once selection biases were accounted for we would find that the timing of reproductive and cardio-metabolic health for Black and Hispanic women has been underestimated.

Aim 1 created a framework for examining selection bias in a cohort and quantified the extent to which selection at study commencement may bias racial/ethnic differences in SWAN. Hypotheses included:

H 1.1 Black and Hispanic women would have the lowest rates of selection into the SWAN cohort compared to other racial/ethnic groups.

H 1.2 Black and Hispanic women would be subject to higher levels of left truncation than other racial/ethnic groups due to early natural reproductive aging and surgical menopause.

Aim 2 corrected for selection bias, specifically left truncation and right censoring and estimated their effects on racial/ethnic differences on a main outcome of the SWAN study, the timing of the final menstrual period (FMP). Hypotheses included:

H 2.1 Contrary to past SWAN longitudinal papers[17] that reported no racial/ethnic differences in age at FMP (after adjustment), correcting for selection bias will result in significant racial/ethnic differences in age of FMP.

H 2.2 Given preliminary evidence to suggest earlier reproductive aging[17–20] and evidence showing higher reproductive morbidity[20–23], after correcting for selection bias Black and Hispanic women will have earlier reproductive aging compared to White women.

Aim 3 corrected for selection bias, specifically left truncation, left censoring and right censoring, and estimated their effects on racial/ethnic differences in the timing of various secondary cardio-metabolic outcomes in the SWAN study, including hypertension, isolated systolic hypertension, insulin resistance and diabetes. The hypothesis was:

H 3.1 Given evidence of weathering [24–27] and the high prevalence of cardio-metabolic mortality and morbidity among Black and Hispanic women[28,29], Black and Hispanic women will have earlier onset of hypertension, isolated systolic

hypertension, insulin resistance, and diabetes as compared to White women after accounting for selection.

1.2 Background

Many epidemiologic studies are geared toward examining disparities in the prevalence of an outcome of interest[4,5,30–35]. However, racial disparities can also be conceptualized from a life-course perspective[36–38] as the differences in the timing, or average age of onset, of a particular outcome. Various studies have found that cardio-metabolic risk factors and outcomes, such as hypertension, diabetes, stroke, and cardiovascular disease, tend to have a higher prevalence at earlier ages in Blacks and Hispanics compared to Whites[5,6,8,9]. Earlier occurrence of markers of aging and major diseases can lead to earlier disability and death, leading to well-documented and persistent racial/ethnic differences in lifespan[24–26,29]. Given that a typical cohort is assembled with the underlying assumption that "normative" aging is what occurs in White populations, the differences in typical ages of onset could cause selection bias, or a distortion in estimation caused by differential selection into or out of a study or analyses[2,39]. Extending established evidence of racial differences in cardio-metabolic health and mixed evidence of racial differences in reproductive aging, this dissertation estimates the degree to which selection bias affects racial/ethnic disparities in the estimated timing of final menstrual period (FMP) and cardio-metabolic outcomes in SWAN.

Racial Differences in Cardio-Metabolic Health Declines. Life expectancy for Black Americans is approximately 3.8 years lower than their White counterparts [29]. For Black women the life expectancy difference is 3.3 years lower than White women, mostly due to higher rates of cardio-metabolic diseases such as heart disease, stroke and diabetes [29], with the largest disparity occurring at midlife and early old age [24–27]. Literature to date supports the hypothesis that earlier aging and earlier onset of disease leading to decreased life expectancy may be due to "weathering" or early health deterioration as a consequence of the cumulative impact of repeated experience with social or economic adversity and political marginalization [6]. In support of this theory, combined measures of cardio-metabolic risk factors (such as c-reactive protein, serum creatinine, systolic blood pressure, total cholesterol and body mass index) have been shown to increase with age at a higher rate in Black women than in White women[6,8,9]. Systolic and

diastolic blood pressure as individual markers have been shown to increase with age at faster rates in Black and Hispanic versus White women, with diastolic blood pressure peaking earlier and then starting to decrease (upside down u-shape) earlier in Black and Hispanic versus White women[40–42]. Little research has shown an earlier onset for racial/ethnic minorities versus White's for insulin resistance and diabetes, however younger prospective cohorts have noted consistently higher overall incidence of these metabolic conditions among racial/ethnic minorities versus Whites[43–45]. And the prevalence of various cardio-metabolic conditions such as hypertension and diabetes has been found to be greater at younger ages in Black versus White women further suggesting accelerated health declines[5]. Few studies, to our knowledge, have examined racial/ethnic differences in the average age of onset of individual cardio-metabolic outcomes in a longitudinal cohort.

Racial Differences in the Timing of Menopause. Reproductive aging has been posited as a vital marker for women's health, particularly for cardio-metabolic health in the midlife period when women are at highest cardio-metabolic risk [46–49]. Reproductive aging is signaled by a women's final menstrual period (FMP)[46–49]. The occurrence and timing of the FMP has been hypothesized to be a potent cardio-metabolic risk factor, as an earlier age at FMP has been associated with increased risk of cardiovascular disease, mortality from ischemic heart disease, and stroke [49–51] as well as with risk factors such as low HDL, increased waist circumference, and hypertension [52]. Earlier menopause is hypothesized to increase risk by prolonging the time that a woman experiences low reproductive steroid hormone production[53] potentially causing earlier onset of cardio-metabolic disorders. Some women reach their FMP naturally while others have a hysterectomy or bilateral oopherectomy which subsequently ceases menstrual bleeding. Previous studies have highlighted the high prevalence of surgical menopause among Black and Hispanic women, especially Black women[54–56]. The differential risk of hysterectomy/oophorectomy could be partially due to the higher prevalence of reproductive morbidities (such as uterine fibroids) earlier in life for Black women, which are often treated with surgery[57–59]. However, overall the evidence is mixed regarding whether or not racial/ethnic differences exist in the timing of natural FMP[17–23,60] and to our knowledge, there is little longitudinal evidence of racial/ethnic differences in the typical age of surgical FMP[55,61–65].

Selection Bias and Left Truncation. The lack of consistent findings regarding racial/ethnic differences in age at the FMP could be due partially to differences in the extent of selection bias across studies. Furthermore, racial differences in cardio-metabolic outcomes could be underestimated due to selection bias. Selection bias arises from differential probability of selection of participants into and out of a study sample and may distort study estimates as well as limit the generalizability of estimates to the population of interest [2]. Cohort studies are particularly subject to such bias via selection into (left truncation and left censoring) and out of (right censoring) the cohort. Left truncation occurs when individuals who have already had the outcome of interest are not included in the study. Left censoring occurs when individuals who have had the outcome of interest are included in the study. And right censoring occurs when the outcome of interest is not observed for a participant in a cohort study[3], either from non-occurrence of the outcome prior to study termination, from loss to follow up, or due to missing variables needed to determine the outcome of interest.

To illustrate the forms of selection present in a cohort, consider an example of how participants with different profiles were selected into and out of the longitudinal Study of Women's Health Across the Nation (SWAN). The SWAN cohort recruited women between the ages of 42 to 52 years. Women who had their FMP prior to recruitment, a hysterectomy/oophorectomy, or were taking hormones at the time of recruitment were not selected into the cohort (left-truncated).

In Figure 1 below we represent two women that according to their ages would have been eligible for the cohort but were not included due to left truncation, one woman that was included into the cohort but had HTN before entry into the cohort (left censored) and three women selected out of the cohort or right censored.

Figure 1.1 Examples of Left Truncation, Left Censoring and Right Censoring in SWAN

	Years from Baseline (BL)	-5	-4	-3	-2	-1	BL	1	2	3	4	5	6	7	8	9	10	11	12	13
	Left Truncated Due to:																			
	1. Age of FMP	37	38	39	FMP															
	2. Study Timing	46	47	48	FMP															
	3. Selection Criteria	37	38	Surgery																
	Observed	37	38	39	40	41	42	43	44	45	46	47	48	49	FMP	51	52	HTN	54	55
Age	Left Censored for HTN:																			
	4. HTN at Baseline	40	HTN	42	43	44	45	46	47	48	49	FMP								
	Right Censored Due to:																			
	5. Loss to Follow Up	40	41	42	43	44	45	46	47	48	49	Dropped Out								
	6. Surgical Amenorrhea	39	40	41	42	43	44	45	46	47	48	49	50	51	52	Surgery				
	7. Non-Occurrence of Outcome	46	48	47	48	49	50	51	FMP	53	54	55	56	57	58	59	60	61	62	63

Note Eligibility criteria: 42-52 years of age, no reproductive surgery, no hormone use and no FMP

For the estimation of the timing of FMP, left truncation and left censoring are relevant forms of selection, corresponding to women 1-3 and women 6-7 respectively. No left censoring will occur because no women made it into the SWAN cohort who had already had their FMP due to the eligibility criteria[15]. Women that were ineligible for the cohort were left truncated, left truncation can happen for many reasons (women 1-3). For example, woman 1 had an early FMP (age 40) and was subsequently ineligible for the cohort. Woman 2 had "normal" FMP timing[49] at age 49, but at the time of enrollment she was age eligible (age 51) but not premenopausal, thus not eligible for the cohort. Woman 3 was age eligible (Age 43) but had a hysterectomy at age 40, making her ineligible. Women that were eligible and included in the cohort but were lost to follow up were right censored (women 5-6). For example, Woman 5 dropped out of the study before an outcome was observed. In previous SWAN analyses and other studies of the natural FMP[17,20–23,54,66], women who are eligible and included in the cohort but had a surgical FMP in follow up are also right censored corresponding to woman 6 who had a hysterectomy at age 51 before her FMP thus was right censored for the estimation of FMP.

For the estimation of the timing of cardio-metabolic outcomes, in this example the timing of hypertension (HTN), left truncation, left censoring and right censoring are relevant forms of selection. Left truncation would occur for the same reasons, i.e., women that were ineligible for the cohort due to the FMP eligibility criteria would be left truncated (women 1-3). However, estimation of a secondary outcome introduces the potential for left censoring, i.e., women that were eligible and included in the cohort but had the outcome (HTN) prior to entry into to study or at baseline. This scenario would be like woman 4 who was left censored because she had onset of HTN before her entry into the cohort. Lastly, women that were eligible and included in the cohort but lost to follow up are right censored (women 5 and 7). Woman 5 was lost to follow up before the observation of any outcomes. Woman 7 would be included for the estimation of FMP but would be right censored for HTN, as she did not develop HTN during the study.

All three forms of selection may occur in any cohort; however, they can cause selection bias if the probabilities of left truncation, left censoring or right censoring *differ* by population sub-groups of interest. In the case of left truncation, if the recruitment age differentially misses subgroups of the population that experience the outcome earlier in life, the study may

overestimate the average age of onset and incidence of the disease or outcome of interest and underestimate racial/ethnic disparities. Some studies have found evidence of earlier menopause in Black women[18,21,23,60]. In one prospective study, Black women had their FMP one year earlier on average than White women (49.3 versus 51.5 years). This disparity increased for Black women who reported psychosocial stress (48.4 years)[60]. The SWAN cross-sectional screening survey (age range 40-55 years) has reported that Black women were more likely than White women to experience premature ovarian failure (cessation of menstruation before the age of 40), but did not show racial differences in natural menopausal timing except that Japanese women had later natural menopausal timing compared to other racial/ethnic groups[20]. The SWAN longitudinal cohort (age range 42-52 years) reported no Black/White differences in age of onset of the natural FMP in adjusted analyses[17]. However, it is possible that the SWAN longitudinal cohort differentially selected out women with earlier menopause by race/ethnicity, thereby biasing estimates of racial/ethnic differences in FMP timing. Left truncation can also be caused by a specific exclusion criterion that differentially affects subgroups of the population. For example, in SWAN approximately 31% of Black women versus 17% of White women had surgical amenorrhea (loss of menses due to hysterectomy or oophorectomy)[20] at the time of enrollment, precluding these women from entering the cohort. Overall, Black women have a higher rate of surgical menopause, due to hysterectomies or oophorectomies, than other racial/ethnic groups and thus are frequently excluded from analyses of timing of menopause [17,20–23,66]. The differential risk of hysterectomy/oophorectomy could be partially due to the higher prevalence of reproductive morbidities (such as uterine fibroids) earlier in life for Black women, which are often treated with surgery [57–59]. Accordingly, in contrast to some other studies[18,21,23,60], the SWAN longitudinal study has not found evidence of Black/White differences in the timing of the FMP other than in differences in premature ovarian failure in the cross sectional screening survey[21]. It is possible that this negative finding is at least partially due to selection bias issues stemming from differential left-truncation in the SWAN cohort which may affect the estimation of other outcomes as well.

1.3 Analytic Approach

Study Design and Population. Data were from the Study of Women's Health Across the Nation (SWAN), a 25-year multi-site, multi-racial/ethnic cohort designed to study women as they

underwent the menopausal transition and the impact of menopause on women's risk of chronic disease. The SWAN study includes 7 clinical sites, each enrolled White women, and women from another race/ethnic group (Black women at 4 sites; Chinese, Japanese, and Hispanic women at one site each). White women were recruited all seven study sites; Black women at Chicago, Detroit, Boston, and Pittsburgh; Hispanic women at New Jersey; Chinese women in Oakland; and Japanese women in Los Angeles. Data were used from 2 stages of the SWAN cohort: 1) the cross-sectional screening survey of 16,605 women aged 40 to 55 years of age used to identify eligible women, and 2) the 3,302 women enrolled into the SWAN cohort [15]. The cross sectional screening survey recruited via established sampling frames yielding 16,605 women who were 40-55 years old, self-identified as the racial/ethnic group of the site and were able to give verbal consent to participate[15]. Subsequently, 3302 eligible women (aged 42-52 years, premenopausal with an intact uterus and at least one ovary, and not using hormone therapy[15]) were enrolled into the SWAN longitudinal cohort. The baseline, and the 15 approximately annual follow-up visits provided detailed information regarding menopausal timing, socio-demographic factors, lifestyle, and blood samples for cardio-metabolic biomarkers as women transitioned from pre- to post-menopause. The institutional review board of each study site approved the protocol and women provided informed consent prior to each visit. At the 15th follow-up, 149 women were deceased, and SWAN remained in contact with 2,333 (74%) of the surviving women.

Final Menstrual Period (FMP). Age at FMP for cohort participants was based on detailed interview data regarding bleeding patterns and reproductive history. Natural FMP status was determined for each woman at each clinic visit with the date of the FMP being the last reported bleeding date after 12 months of amenorrhea were observed, not attributable to pregnancy/breastfeeding. Additionally, women annually reported if they used hormones in the past year or underwent a bilateral oophorectomy/hysterectomy. For the purposes of this analysis, age of FMP was the recorded age at which a woman either reports 12 months of amenorrhea, or the date of surgery for a bilateral oophorectomy/hysterectomy. Of the 3,302 women included in the longitudinal cohort 57.6% had an observed natural FMP, 5.7% have an unobserved FMP due to surgery, 36.7% are unobserved due to loss to follow up or masking by hormone use. FMP

status (self-reported age at FMP or surgical amenorrhea) was also available in the screening survey and was used for the inverse probability weighting analysis in Aim 1.

Cardio-Metabolic Outcomes. The main cardio-metabolic outcomes of interest were hypertension, isolated systolic hypertension, insulin resistance and diabetes. At each clinic visit systolic and diastolic blood pressure, hypertension medications, insulin sensitivity with the homeostatic model assessment for insulin resistance (HOMA-IR), fasting blood glucose measurements, diabetes medication and physician diagnosed diabetes was collected. Hypertension was defined as systolic or diastolic blood pressure greater than or equal to 140/90 mmHg or on blood pressure medication while isolated systolic hypertension was defined as systolic blood pressure greater than or equal to 140 mmHg with a diastolic blood pressure of less than 90 mmHg (hypertension medications were used as a covariate for isolated systolic blood pressure)[67]. Insulin resistance was defined as having a HOMA-IR value > 5.9 (based on previously established cut-points [68,69]) or taking insulin. Diabetes was defined as reporting a physician diagnosis of diabetes, fasting glucose > 126 mg/dL and/or use of diabetes medication. For the purposes of this analysis, the outcome was age at which a woman becomes either hypertensive, isolated systolic hypertensive, insulin resistant or diabetic.

Demographics and Covariates. Racial/ethnic group was self-identified as White, Black, Hispanic, Chinese, or Japanese. Covariates present in both the cross-sectional screening and the longitudinal cohort study included hormone use, demographic information such as socioeconomic status (education, income, financial stress, neighborhood factors), health behaviors (physical activity, energy intake, smoking, alcohol, and drug use), full medical history, and past reproductive medical history (e.g., age of menarche, parity, birth outcomes, uterine fibroids).

Analytic Approach. Aim 1 exclusively used SWAN's initial cross-sectional screening study to examine selection mechanisms and assess the potential for left truncation. First, sampling methodologies were reviewed to identify selection mechanisms.

Next, to quantify how each mechanism affected the racial/ethnic distribution within the cohort, frequencies and percentages of women who did and did not make it through each stage of selection were calculated for the total sample and by racial/ethnic group. Then to quantify the potential bias in estimating racial/ethnic disparities related to timing of menopause and chronic disease outcomes, the association between each type of selection identified and the following predictors were estimated: racial/ethnic group (exposure); common causes (Z) of the main outcomes of the study; and moderation by racial/ethnic group on common causes (Z).

Associations were estimated using bivariate analyses and multivariable logistic regression. Multivariable imputation by chained equations (MICE)[70] was used to estimate missing variable values for women retained in the cohort. Demographic, reproductive, and health predictors from the cross-sectional screening and baseline visit of the cohort study were used to predict the missing variables. Forward and backward selection using the Akaike Information Criterion (AIC) was implemented to select predictors for the multi-variate models.

Aim 2 utilized inverse probability weighting and multiple imputation to assess the impact of selection bias and to account for bias in subsequent analyses. Inverse probability weighting (IPW) is a statistical procedure that applies weights to observations in a cohort that are representative of persons who were excluded, either because of biased selection into or out of the cohort, thereby creating an unbiased pseudo-population from which to draw inferences[71–73]. Multiple imputation (MI) is a statistical procedure that replaces missing data based on an imputation model, which is a model of the missing data given the observed data. The missing data are imputed repeatedly over a specified number of datasets with the final estimates pooled across the datasets[71]. IPW is ideal for accounting for selection bias incurred due to differential selection *into* the cohort, as it simply weights the retained cohort participants data based on probability of selection, rather than assuming that an imputation model can correctly impute the unobserved longitudinal data for each underrepresented participant. MI is ideal for accounting for selection bias incurred due to selection *out* of the cohort, or loss to follow up, assuming the data is missing at random, as it utilizes collected data to individually impute subsequent missing data points, maximizing statistical efficiency.

SWAN has explored missing data patterns and has recently multiply imputed[70] age at natural FMP for all women who were lost to follow up, started hormones before FMP or underwent a hysterectomy or oophorectomy based on demographic, health behavior, medical history and reproductive history. The current analysis modified this imputation to include surgical FMP. Thus, the imputation was undertaken in two steps. The first step did not assume that all person's loss to follow up had a natural FMP but imputed the type of FMP for each person with missing information. The second step used the FMP type as an additional predictor for the age of FMP. Then, using the imputed data, IPW was applied to correct for the inherent design of the study that selected women aged 42-52 years who hadn't undergone menopause, calculated as the probability of a woman aged 42 to 52 being excluded from the study given her age and menopausal status, which is proportional to the time she was eligible (premenopausal) over the 10-year period between 42 and 52 years. The inverse of this probability was used to weight the cohort and make it exchangeable to the sample that would have been retained had all women been 42 years of age upon recruitment.

The second IPWs [71–73] used the imputed dataset to up-weight cohort participants that are similar to women not included in the cohort from the cross-sectional screening study due to eligibility (eligibility weight) and among those eligible, non-participation (participation weight). Weights were calculated as the reciprocal probability of being included in the SWAN cohort study from the cross-sectional screening based on cross sectional screening traits such as racial/ethnic group, demographics (such as socioeconomic status) and past reproductive and medical history.

Lastly, we multiplied the three IPW weights and used imputed data to assess the impact of simultaneously accounting for selection biases on estimates of average age of FMP and racial/ethnic differences in average age of FMP. Gold et al. published an analysis of average age of FMP not accounting for potential selection bias[17]. Like Gold et al, Cox Proportional Hazard models were used to estimate FMP and race specific estimates of FMP with age as the timescale. Models controlled for covariates: smoking, self-reported health at baseline, educational level at baseline, baseline use of oral contraceptives, time varying alcohol use, time-varying employment, time-varying physical activity score and weight. Race/ethnic specific age at FMP

estimates, for each type of FMP (natural and surgical), were compared across the models below to assess the impact of each source of selection bias.

Aim 3 analyses corrected for left truncation, left censoring and right censoring, and estimated the racial/ethnic differences in timing of onset of four cardio-metabolic outcomes (hypertension, isolated systolic hypertension, insulin resistance and diabetes). Left truncation was addressed using the IPW weights developed in aim 2. Right censoring was addressed using multiple imputation. And left censoring was addressed by using accelerated failure time (AFT) models with age as the timescale, a flexible model that accounts for prevalent cases at baseline[3]. Multiple imputation was again undertaken in 2 steps. First to determine if the women who were lost to follow up had the outcome of interest during the course of the study. Then within each of the imputed datasets occurrence of the outcome was used to subsequently predict the age at onset of the outcome, allowing the age of onset to vary according to whether the outcome had occurred. AFT models were estimated for each of the cardio-metabolic risk factors to determine if there are racial/ethnic differences in age of cardio-metabolic risk onset. Each model was estimated with correction for different forms of selection bias to examine the impact of each bias on the association between racial/ethnic group and the outcome in the following order: unadjusted for selection, adjusted for left censoring, last model with additional adjustment for right censoring, last model with additional adjustment for left truncation.

1.4 Summary

As selection bias potentially affects racial/ethnic groups differentially, this work develops a framework and approach to account for selection bias that may be adapted and applied to many different cohort studies interested in disparities in aging. Results of this research contribute to research on aging by using IPW weights in a novel way to address left truncation, an understudied source of selection bias in cohorts. Aims 2 and 3 demonstrate the impact of simultaneously correcting for selection biases on primary and secondary outcomes in SWAN, information that is critical to correctly estimating average age of disease onset and disparities in aging. Lastly, results of this research have the potential to add to the body of research on "weathering" by examining the extent of racial/ethnic differences in accelerated health declines across reproductive and cardio-metabolic health outcomes.

Chapter 2 of this dissertation presents the analysis and results for Aim 1 which examines the potential effects of selection bias on racial/ethnic differences in FMP and related outcomes in SWAN. Chapter 3 presents the analysis and results for Aim 2 which corrects for left truncation and right censoring and estimates their effects on racial/ethnic differences in timing of FMP, or age at menopause. Chapter 4 presents the analysis and results for Aim 3 which corrects for left truncation, left and right censoring in the estimation of racial/ethnic differences in the age of cardio-metabolic outcome onset. Lastly, chapter 5 provides additional context for the intuition behind this dissertation, reviews the strengths and limitations, the public health implications of the findings and future research directions for the dissertation as whole.

1.5 References

1. Doll, R. Cohort studies: History of the method I. Prospective cohort studies. *Soz. Praventivmed.* **46**, 75–86 (2001).
2. Rothman, K. J., Greenland, S. & Lash, T. J. *Modern Epidemiology*. (Lippincott Williams and Wilkins, 2014).
3. Cain, K. C. *et al.* Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes. *Am. J. Epidemiol.* **173**, 1078–1084 (2011).
4. Geronimus, A. T., Bound, J., Keene, D. & Hicken, M. Black-White Differences in Age Trajectories of Hypertension Prevalence Among Adult Women and Men, 1999-2002. *Ethn. Dis.* **17**, 40–48 (2007).
5. Thorpe, R. J. *et al.* Accelerated Health Declines among African Americans in the USA. *J. Urban Heal.* **93**, 808–819 (2016).
6. Geronimus, A. T., Hicken, M., Keene, D. & Bound, J. 'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am. J. Public Health* **96**, 826–833 (2006).
7. Forrester, S. *et al.* Racial differences in weathering and its associations with psychosocial stress: The CARDIA study. *SSM - Popul. Heal.* **7**, 100319 (2019).
8. Levine, M. E. & Crimmins, E. M. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc. Sci. Med.* **118**, 27–32 (2014).
9. Geronimus, A. T. *et al.* Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-based Sample. *J. Health Soc. Behav.* **56**, 199–224 (2015).
10. Howard, V. J. *et al.* The reasons for geographic and racial differences in stroke study: Objectives and design. *Neuroepidemiology* **25**, 135–143 (2005).
11. Hill, C. *et al.* The Atherosclerosis Risk in Communities (ARIC) Study: Design and Objectives. **129**, (1989).
12. Sonnega, A. *et al.* Cohort profile: The Health and Retirement Study (HRS). *Int. J. Epidemiol.* **43**, 576–585 (2014).
13. Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: Objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
14. Friedman, G. D. *et al.* Cardia: study design, recruitment, and some characteristics of the examined subjects. *J. Clin. Epidemiol.* **41**, 1105–1116 (1988).
15. Sowers, M. *et al.* SWAN: A Multicenter, Multiethnic, Community Based Cohort Study of Women and the Menopausal Transition. in *Menopause: Biology and Pathobiology* 175–188 (2000).
16. Ferrell, R. J. & Sowers, M. Longitudinal, epidemiologic studies of female reproductive aging. *Ann. N. Y. Acad. Sci.* **1204**, 188–197 (2010).
17. Gold, E. B. *et al.* Factors related to age at natural menopause: Longitudinal analyses from SWAN. *Am. J. Epidemiol.* (2013). doi:10.1093/aje/kws421
18. Butts, S. F. & Seifer, D. B. Racial and ethnic differences in reproductive potential across the life cycle. *Fertil. Steril.* **93**, 681–690 (2010).
19. Delellis Henderson, K., Bernstein, L., Henderson, B., Kolonel, L. & Pike, M. C. Original Contribution Predictors of the Timing of Natural Menopause in the Multiethnic Cohort Study. *Am. J. Epidemiol.* **167**, 1287–1294 (2008).
20. Gold, E. B. *et al.* Factors Associated with Age at Natural Menopause in a Multiethnic Sample of Midlife Women. *Am. J. Epidemiol.* **153**, 865–874 (2001).

21. Luborsky, J. L., Meyer, P., Sowers, M. F., Gold, E. B. & Santoro, N. Premature menopause in a multi-ethnic population study of the menopause transition*. *Hum. Reprod.* **18**, 199–206 (2002).
22. MacMahon, B. & Worcester, J. *Age at menopause. United States--1960-1962. Vital and health statistics. Series 11, Data from the National Health Survey* (1966).
23. Stanford, J. L., Hartge, P., Brinton, L. A., Hoover', R. N. & Brookmeyer, R. Factors Influencing the Age of Natural Menopause. *J. Chronic Dis.* **40**, 995–1002 (1987).
24. Nan, |, Astone, M., Ensminger, M. & Juon, H. S. *Early Adult Characteristics and Mortality Among Inner-City African-American Women*. *American Journal of Public Health* **92**, (2002).
25. Geronimus, A. T., Bound, J., Waidmann, T. A., Hillemeier, M. M. & Burns, P. B. Excess Mortality among Blacks and Whites in the United States. *N. Engl. J. Med.* **335**, 1552–1558 (1996).
26. Geronimus, A. T., Bound, J., Waidmann, T. A., Colen, C. G. & Steffick, D. Inequality in Life Expectancy, Funcational Status, and Active Life Expectancy Across Selected Black and White Populations in the United States. *Demography* **38**, 227–251 (2001).
27. Elo, I. T. & Preston, S. H. *Educational Differential in Mortality: United States*. *Soc. Sci. Med* **42**, (1996).
28. Wong, M. D., Shapiro, M. F., Boscardin, W. J. & Ettner, S. L. Contribution of Major Diseases to Disparities in Mortality. *N. Engl. J. Med.* **347**, 1585–1592 (2002).
29. Kochanek, K. D., Arias, E. & Anderson, R. N. *How Did Cause of Death Contribute to Racial Differences in Life Expectancy in the United States in 2010? NCHS data brief, no 125* (2013).
30. Singh, G. K., Siahpush, M., Azuine, R. E. & Williams, S. D. Widening Socioeconomic and Racial Disparities in Cardiovascular Disease Mortality in the United States ,. **3**, 106–118 (2015).
31. Bell, C. N., Thorpe, R. J., Bowie, J. V. & LaVeist, T. A. Race disparities in cardiovascular disease risk factors within socioeconomic status strata. *Ann. Epidemiol.* **28**, 147–152 (2018).
32. Fuller-Rowell, T. E. *et al.* Disparities in insulin resistance between black and white adults in the United States: The role of lifespan stress exposure. *Psychoneuroendocrinology* **107**, 1–8 (2019).
33. Hasson, B. R., Apovian, C. & Istfan, N. Racial/Ethnic Differences in Insulin Resistance and Beta Cell Function: Relationship to Racial Disparities in Type 2 Diabetes among African Americans versus Caucasians. *Curr. Obes. Rep.* **4**, 241–249 (2015).
34. Lin, S. X., Carnethon, M., Szklo, M. & Bertoni, A. Racial/ethnic differences in the association of triglycerides with other metabolic syndrome components: The multi-ethnic study of atherosclerosis. *Metab. Syndr. Relat. Disord.* **9**, 35–40 (2011).
35. Falkner, B. & Cossrow, N. D. F. H. Prevalence of metabolic syndrome and obesity-associated hypertension in the racial ethnic minorities of the United States. *Curr. Hypertens. Rep.* **16**, (2014).
36. Gee, G. C., Hing, A., Mohammed, S., Tabor, D. C. & Williams, D. R. Racism and the Life Course: Taking Time Seriously. *Am J Public Heal.* **109**, 43–47 (2019).
37. Ben-Shlomo, Y. & Kuh, D. A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *Int. J. Epidemiol.* **31**, 285–293 (2002).

38. Braveman, P. What is health equity: And how does a life-course approach take us further toward it? *Matern. Child Health J.* **18**, 366–372 (2014).
39. Hernan, M. A. & Robins, J. M. *Causal Inference*. (2016).
40. Chobanian, A. V *et al.* Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatement of High Blood Pressure. *Hypertension* **42**, 1206–1252 (2003).
41. Burt, V. L. *et al.* Prevalence of Hypertension in the US Adult Population. *Hypertension* **25**, 305–313 (1995).
42. Zeki, A. *et al.* Use of a pooled cohort to impute cardiovascular disease risk factors across the adult life course. *Int. J. Epidemiol.* **0**, (2018).
43. Lee, J. M., Okumura, M. J., Davis, M. M., Herman, W. H. & Gurney, J. G. Prevalence and Determinants of Insulin Resistance Among U.S. Adolescents A population-based study. (2006). doi:10.2337/dc06-0709
44. Brancati, F. L., Kao, W. H. L., Folsom, A. R., Watson, R. L. & Szklo, M. Incident Type 2 Diabetes Mellitus in African American and White Adults. *JAMA* **283**, 2253 (2000).
45. Ford, E. S. Prevalence of the Metabolic Syndrome Defined by the International Diabetes Federation Among Adults in the U.S. *Diabetes Care* **28**, 2745–2749 (2005).
46. Snowdon, D. A. *et al.* Is Early Natural Menopause a Biologic Marker of Health and Aging? *Am. J. Public Health* **79**, 709–714 (1989).
47. Wise, P. M., Krajnak, K. M. & Kashon, M. L. Menopause: The Aging of Multiple Pacemakers. *Am. Assoc. Adv. Sci.* **273**, 67–70 (1996).
48. Cooper, G. S. & Sandler, D. P. Age at Natural Menopause and Mortality. *Ann. Epidemiol.* **8**, 229–235 (1998).
49. Gold, E. B. The timing of the age at which natural menopause occurs. *Obstet. Gynecol. Clin. North Am.* **38**, 425–40 (2011).
50. Appelman, Y., van Rijn, B. B., ten Haaf, M. E., Boersma, E. & Peters, S. A. E. Sex differences in cardiovascular risk factors and disease prevention. *Atherosclerosis* **241**, 211–218 (2015).
51. Muka, T. *et al.* Association of Age at Onset of Menopause and Time Since Onset of Menopause With Cardiovascular Outcomes, Intermediate Vascular Traits, and All-Cause Mortality. *JAMA Cardiol.* **1**, 767 (2016).
52. Bleil, M. E., Gregorich, S. E., Mcconnell, D., Rosen, M. P. & Cedars, M. I. DOES ACCELERATED REPRODUCTIVE AGING UNDERLIE PRE-MENOPAUSAL RISK FOR CARDIOVASCULAR DISEASE? *Menopause* **20**, 1139–1146 (2013).
53. Mendelsohn, M. E. & Karas, R. H. The Protective Effects of Estrogen on the Cardiovascular System. *N. Engl. J. Med.* **340**, 1801–1811 (1999).
54. WRIGHT, M. A. *et al.* Changing trends in Black-White racial differences in surgical menopause: a population-based study. *Am. J. Obstet. Gynecol.* (2021). doi:10.1016/j.ajog.2021.05.045
55. Powell, L. H. *et al.* Ethnic differences in past hysterectomy for benign conditions. *Women's Heal. Issues* **15**, 179–186 (2005).
56. Bower, J. K., Schreiner, P. J., Sternfeld, B. & Lewis, C. E. Black-white differences in hysterectomy prevalence: The CARDIA study. *Am. J. Public Health* **99**, 300–307 (2009).
57. Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. High cumulative incidence of uterine leiomyoma in black and white women: Ultrasound evidence Donna. *Am. J. Obstet. Gynecol.* **188**, 100–107 (2003).

58. Marsh, E. E. *et al.* Racial differences in fibroid prevalence and ultrasound findings in asymptomatic young women (18-30 years old): a pilot study. *Fertil. Steril.* **99**, 1951–1957 (2013).
59. Okolo, S., Medical Executive Director, F. & of Ob-Gyn, P. Incidence, aetiology and epidemiology of uterine fibroids. *Best Pract. Res. Clin. Obstet. Gynaecol.* **22**, 571–588 (2008).
60. Bromberger, J. T. *et al.* Prospective Study of the Determinants of Age at Menopause. *Am. J. Epidemiol. Copyr. O* **145**, (1997).
61. Kjerulff, K. H. *et al.* Hysterectomy and Race. *Obstet. Gynecol.* **82**, 757–764 (1993).
62. Brett, K. M., Marsh, J. V. R. & Madans, J. H. Epidemiology of hysterectomy in the United States: Demographic and reproductive factors in a nationally representative sample. *J. Women's Heal.* **6**, 309–316 (1997).
63. Marshall, L. M. *et al.* Variation in the incidence of uterine leiomyoma among premenopausal women by age and race. *Obstet. Gynecol.* **90**, 967–973 (1997).
64. Palmer, J. R., Rosenberg, L., Wise, L. A., Horton, N. J. & Adams-Campbell, L. L. Onset of Natural Menopause in African American Women. *Am. J. Public Health* **93**, 16–18 (2003).
65. Palmer, J. R., Rao, R. S., Adams-Campbell, L. L. & Rosenberg, L. Correlates of hysterectomy among African-American women. *Am. J. Epidemiol.* **150**, 1309–1315 (1999).
66. Merrill, R. M., Layman, A. B., Oderda, G. & Asche, C. Risk Estimates of Hysterectomy and Selected Conditions Commonly Treated with Hysterectomy. *Ann. Epidemiol.* **18**, 253–260 (2008).
67. Guidelines, A. C. of C. H. A. T. F. on C. P. *2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults*. *Journal of the American College of Cardiology* (2017). doi:10.1016/j.jacc.2017.11.006
68. Gayoso-Diz, P. *et al.* Insulin resistance (HOMA-IR) cut-off values and the metabolic syndrome in a general adult population: effect of gender and age: EPIRCE cross-sectional study. *BMC Endocr. Disord.* **13**, 47 (2013).
69. Bardini, G., Barbaro, V., Romano, D., Rotella, C. M. & Giannini, S. Different distribution of phenotypes and glucose tolerance categories associated with two alternative proposed cutoffs of insulin resistance. *Acta Diabetol.* **51**, 321–324 (2014).
70. Rubin, D. B. Multiple Imputations in Sample Surveys - A Phenomenological Baysian Approach to Nonresponse. *Proc. Surv. Res. methods Sect. Am. Stat. Assoc. Am. Stat. Assoc.* **1**, 20–34 (1978).
71. Seaman, S. R., White, I. R., Copas, A. J. & Li, L. Combining Multiple Imputation and Inverse-Probability Weighting. *Biometrics* (2012). doi:10.1111/j.1541-0420.2011.01666.x
72. Seaman, S. R. & White, I. R. Review of inverse probability weighting for dealing with missing data. doi:10.1177/0962280210395740
73. Cole, S. R. & Hernan, M. A. Constructing Inverse Probability Weights for Marginal Structural Models. *Am. J. Epidemiol.* **168**, 656–664 (2008).

Chapter 2 Left Truncation and Selection Bias in a Cohort Study of Midlife Aging: Implications for Estimating Racial/Ethnic Disparities

2.1 Introduction

Left truncation occurs when participants who have already experienced the outcome of interest are not included in a cohort study[1,2]. Although selection bias due to loss to follow up, survival and right censoring in cohort studies has been well studied [3–5], less work has focused on selection at study commencement and left-truncation[1,2,6,7]. As cohort studies aim to observe the occurrence of a particular outcome(s), studies typically sample and recruit study participants before the "expected" age of onset in a population, excluding persons who have already experienced the outcome at study commencement. A typical cohort study would be subject to variable left-truncation, i.e. left-truncation that occurs at 'stochastically different timepoints' for each participant[2], however, variable left truncation can result in bias and a loss of precision in estimates if the probability of left truncation differs across the exposure or confounders of interest[2,6]. Furthermore, selection into a cohort may cause collider-stratification bias, or bias stemming from stratifying on a variable (selection in this case) that is effected by two or more variables that are related to the exposure/outcome process[3,7].

Consideration of left truncation bias is particularly important when estimating racial/ethnic disparities in health, as rates of left truncation in cohorts may vary by racial/ethnic group. Life expectancy for Black Americans is approximately 3.8 years shorter than their White counterparts[8]. For Black women, life expectancy is 3.3 years less than White women, primarily due to higher rates of cardio-metabolic diseases such as heart disease, stroke and diabetes[8], with the largest disparity occurring at midlife and in early old age[9–12]. Literature to date supports the theory that accelerated aging and earlier onset of disease leading to decreased life expectancy may be due to "weathering" or early health deterioration as a consequence of the cumulative

impact of repeated experiences of social or economic adversity and political marginalization[13]. In support of this theory, sets of cardio-metabolic risk factors (body mass index, blood pressure, serum creatinine, total cholesterol, and C-reactive protein) have been shown to begin earlier and/or accelerate faster in Black women than in White women[13–15]. Thus while cohorts typically recruit participants based on meeting inclusion criteria for chronologic age, it is important to consider potential differences in the timing of onset of chronic disease caused by "weathering"[13,16] and how such differences in may lead to differential left truncation by racial group.

This study uses data from the Study of Women's Health Across the Nation (SWAN), a multi-racial, multi-site cohort of midlife aging assembled to observe women as they aged through the menopausal transition. The inclusion criterion for this study result in potential left truncation because women might have had surgical removal of ovaries or hysterectomies, started use of hormones, or reached late perimenopause or full menopause before being age-eligible for the study[17]. We use SWAN's initial cross-sectional screening study to examine selection mechanisms and assess potential for left truncation, quantifying the potential bias in estimating racial/ethnic disparities related to timing of menopause and chronic disease outcomes by estimating the association of 1) racial/ethnic group (exposure), 2) common causes (Z) of the main outcomes of the study, and 3) moderation by racial/ethnic group of common causes (Z), on selection into the SWAN cohort.

2.2 Methods

This analysis used data from the 1995-1997 cross-sectional screening survey (n = 15,695) and resulting longitudinal cohort (n = 3,302) of the Study of Women's Health Across the Nation (SWAN). SWAN has 7 participating sites; Black women were included at 4 sites (Detroit-area, MI; Boston, MA; Chicago, IL; Pittsburgh, PA), Hispanic women at 1 site (Newark, New Jersey), Chinese women at 1 site (Oakland, CA) and Japanese women at 1 site (Los Angeles, CA)[17]. White women were included at all 7 sites[17].

The cross-sectional screening study enrolled community-based samples of eligible women[17,18]. Women were eligible for the screening survey if they were: 40-55 years old, self-identified as the

racial/ethnic group for the designated site, spoke either English or the language for the designated site (Spanish in Newark, Japanese in Los Angeles, Cantonese in Oakland), had a primary residence in the designated geographic area for each site and were able to give informed consent[17,18]. Women were interviewed via phone or in person about their medical and reproductive histories along with demographic/social factors[17,18].

Participants were eligible for the cohort from the cross-sectional screening if they were: 42-52 years old, not currently pregnant, did not have a hysterectomy and/or bilateral oophorectomy, had a menstrual period in the past 3 months (pre to peri menopausal), hadn't used female hormones in the past 3 months (including birth control, fertility drugs, estrogens or progestins, hormone patches or creams, hormone injections or post-menopausal hormones).

A total of 16,096 women were screened and 15,695 women were self-confirmed as matching the designated race/ethnic group for each site. Of these, 6,521 were eligible for the longitudinal cohort and 3,302 women ultimately enrolled. The Institutional Review Boards at each study site approved the protocol and participants provided informed consent for the screening survey and all cohort visits.

Demographic Characteristics. Participants were asked to self-identify their primary racial/ethnic group. The "Black" category included Black, African-American or those identifying as someone of African origin or descent, the Chinese category included those of Chinese origin or Chinese-Americans, the "Hispanic" category included those identifying as Puerto Rican, Mexican/Mexican-American, Dominican, Central American, Cuban/Cuban-American, South American/Spanish/Other Hispanic, the Japanese category included those of Japanese origin or Japanese-Americans, and the "White" category included those identifying as Caucasian. Educational level was obtained by asking about the highest level of education they had attained; responses ranged from did not go to school to doctoral degree (18 categories total). Educational level answers were collapsed into a 3-level category for this analysis (<= High School, Some College, and >= College). Financial hardship was assessed by asking, "How hard is it for you to pay for the *very basics* like food, housing, medical care, and heating?", responses ranged on a 3-point scale from "very hard" to "somewhat hard" to "not very hard at all".

Reproductive and Health Characteristics. Participants indicated if they had ever taken birth control pills for any reason (yes/no) and how many live children they had given birth to (0, 1, 2-3, 4+). Participants rated their general health on a 5-point scale from poor to excellent; the categories poor and fair were collapsed to make a 4-level variable. To calculate body mass index (BMI, kg/m^2), participants estimated their weight in light clothing and their height. For health conditions participants indicated if a doctor, nurse practitioner, or other healthcare provider ever told them that they had: diabetes, heart attack or angina, osteoporosis or brittle/thinning bones, fibroids or benign growths of the uterus or womb, or cancer (other than skin cancer).

Analysis. Sampling methodologies were reviewed to identify selection mechanisms. Next, to quantify how each mechanism affected the racial/ethnic distribution within the cohort, frequencies and percentages of women who did and did not make it through each stage of selection were calculated for the total sample and by racial/ethnic group. Lastly, bivariate analyses and multivariable logistic regression were conducted to estimate the association between hypothesized predictors of selection and the odds of selection. Multivariable imputation by chained equations (MICE)[19] was used to estimate missing variable values for women retained in the cohort. Demographic, reproductive, and health predictors from the cross-sectional screening and baseline visit of the cohort study were used to predict the missing variables. Forward and backward selection using the Akaike Information Criterion (AIC) was implemented to select predictors for the multi-variate models. Analyses were conducted using STATA version 16 and R version 4.0.3.

2.3 Results

Two main selection mechanisms were identified from the cross-sectional screening study: 1) eligibility, meaning the study-design based selection of women from the cross-sectional study who met eligibility criteria for the cohort, and 2) participation, meaning those who were willing and able to participate in the cohort from among the eligible women identified by the cross-sectional study (Figure 1).

Table 1 describes the sample characteristics at both stages of selection. Generally, compared to the cross-sectional screening the women in the cohort were more educated, less likely to have had a diagnosis of fibroids, and in better overall health. The percentage of Black and White participants was relatively stable across both selection stages (Black = 28.1%, 26.2% and 28.3% and White = 49.7%, 49.1%, and 47.0%, respectively) while the proportion of Chinese and Japanese women increased slightly across selection stages (Chinese = 4.2%, 5.6% and 7.5% and Japanese = 5.4%, 6.8% and 8.5%, respectively) and the proportion of Hispanic women decreased across selection stages (12.6%, 12.3% and 8.7% respectively). The proportion of participants at each educational level in the cross-sectional screening and among eligible women was roughly equal (average 33.2%); however, a larger proportion of women with a college level education or greater enrolled into the cohort (42.6% at college level or greater in cohort). Roughly a quarter of the cross-sectional screening sample reported having been diagnosed with fibroids at some point in their lives (25.6%), while just 18.9% in the eligible sample and 20.4% in the enrolled cohort reported a fibroid diagnosis. Roughly 60% of the participants in each stage reported having very good or good health (very good = 31.7%, 34.0% and 35.5% and good = 30.5%, 30.0% and 28.4% respectively). Most participants had a BMI less than 25 kg/m^2 (43.1%, 45.0% and 44.0%, respectively) and had never smoked (52.0%, 55.6%, 58.1%, respectively).

In the total sample (Table 2), 42.0% of screening participants were eligible for the longitudinal cohort, of those eligible 50.65% participated in the cohort. Across racial/ethnic groups, the percent eligible for the cohort was highest among Japanese participants (55.4%) followed by Hispanic (51.9%), White (41.1%), Chinese (40.5%) and then Black (38.9%) participants. Of those eligible for the cohort, the percent who participated in the cohort was highest among Japanese participants (69.1%) followed by Hispanic (63.3%), Black (54.7%), White (48.4%) and Chinese (35.7%) participants. The leading cause of ineligibility for the cohort among the total cross-sectional sample was being outside the age range (30.8%), followed by surgical menopause (20.1%), current hormone use (18.2%), being late-perimenopause or natural post menopause (4.8%) and current pregnancy (0.2%). Black women had an almost 50% higher rate of ineligibility for the cohort due to surgical menopause (30.9%) compared to other groups, which was the leading cause of ineligibility into the cohort for Black women.

In bivariate models for participation in the cohort given eligibility (Table 4), Black, Chinese, and Japanese participants had an increased odds of participation compared to Whites (OR^{Black} = 1.29 (1.14, 1.45); $OR^{Chinese}$ = 2.38 (1.89, 3.01) and $OR^{Japanese}$ = 1.84 (1.50, 2.26)) while Hispanic participants had a decreased odds of participation compared to Whites ($OR^{Hispanic}$ = 0.59 (0.50, 0.69)) (AUC = 0.58). Age was not significantly associated with participation (OR^{age} = 0.99 (0.98, 1.01), AUC = 0.50). Educational level was associated with participation (AUC = 0.59) where lower levels of education were associated with a decreased odds of participation ($OR^{some\ college}$ = 0.76 (0.68, 0.86) and $OR^{<=HS}$ = 0.43 (0.38, 0.48), reference was >= college degree). This association varied by racial/ethnic group (X^2 = 77.64, p-value <0.001; AUC = 0.64), where the trend held for White women ($OR^{somecollege}$ = 0.78 (0.66, 0.92) and $OR^{<=HS}$ = 0.31 (0.26, 0.37)) and was opposite for Hispanic women ($OR^{somecollege}$ = 1.75 (0.84, 3.63) and $OR^{<=HS}$ = 2.25 (1.13, 4.47)). Of the reproductive characteristics, ever using hormonal birth control (OR = 1.41 (1.26, 1.57), AUC = 0.53) and ever being diagnosed with fibroids (OR = 1.21 (1.07, 1.38), AUC = 0.51) was associated with participation. Ever being diagnosed with diabetes, heart attack/angina, osteoporosis, and cancer were not associated with participation. Decreases in level of self-reported health were associated with lower odds of participation compared to an "excellent" health rating (OR^{good} = 0.87 (0.75, 1.00) and $OR^{fair/poor}$ = 0.83 (0.71, 0.98), AUC = 0.52). BMI was associated with participation (AUC = 0.53). Participants with a BMI greater or equal to 30 kg/m^2 had increased odds of participation compared to those with a BMI less than 25 (OR = 1.22 (1.09, 1.37). Smoking status was also associated with participation (AUC = 0.53). Current smokers had a decreased odds of participation compared to participants that had never smoked (OR = 0.67 (0.59, 0.76)).

Model 2 adds an interaction between racial/ethnic group and age (likelihood ratio test: X^2 = 79.05 (p-value <0.001), interaction coefficients = 0: X^2 = 27.90 (p-value <0.001, AUC = 0.70). Black women have the steepest decrease in eligibility to the cohort with increasing age while Chinese and Japanese women have the slowest decrease in eligibility to the cohort with increasing age (OR^{Black} = 0.86 (0.84, 0.89) to $OR^{Chinese}$ = 0.95 (0.90, 1.01)). Figure 2 displays the predicted probabilities of eligibility to the cohort by age at study screening for each racial/ethnic group. Model 3 adds an interaction between racial/ethnic group and educational level (likelihood ratio test: X^2 = 76.29 (p-value <0.001), interaction coefficients = 0: X^2 = 29.43 (p-value <0.001),

AUC = 0.70). White women had decreased odds of eligibility to the cohort with decreasing educational level ($OR^{somecollege}$ = 0.78 (0.69, 0.87) and $OR^{<=HS}$ = 0.65 (0.58, 0.74), reference was >= college degree). Other racial/ethnic groups did not exhibit a significant trend in eligibility by education except Black women with <= HS education (OR = 0.68 (0.49, 0.94)).

In bivariate models for participation in the cohort given eligibility (Table 4), Black, Chinese, and Japanese participants had an increased odds of participation compared to Whites (OR^{Black} = 1.29 (1.14, 1.45); $OR^{Chinese}$ = 2.38 (1.89, 3.01) and $OR^{Japanese}$ = 1.84 (1.50, 2.26)) while Hispanic participants had a decreased odds of participation compared to Whites ($OR^{Hispanic}$ = 0.59 (0.50, 0.69)) (AUC = 0.58). Age was not significantly associated with participation (OR^{age} = 0.99 (0.98, 1.01), AUC = 0.50). Educational level was associated with participation (AUC = 0.59) where decreased in level of education was associated with a decreased odds of participation ($OR^{some\ college}$ = 0.76 (0.68, 0.86) and $OR^{<=HS}$ = 0.43 (0.38, 0.48), reference was >= college degree). This association varied by racial/ethnic group (X^2 = 77.64, p-value <0.001; AUC = 0.64), where the trend held for White women ($OR^{somecollege}$ = 0.78 (0.66, 0.92) and $OR^{<=HS}$ = 0.31 (0.26, 0.37)) and was opposite for Hispanic women ($OR^{somecollege}$ = 1.75 (0.84, 3.63) and $OR^{<=HS}$ = 2.25 (1.13, 4.47)). Of the reproductive characteristics, ever using hormonal birth control (OR = 1.41 (1.26, 1.57), AUC = 0.53) and ever being diagnosed with fibroids (OR = 1.21 (1.07, 1.38), AUC = 0.51) was associated with participation. Ever being diagnosed with diabetes, heart attack/angina, osteoporosis, and cancer were not associated with participation. Decreases in level of self-reported health were associated with lower odds of participation compared to an "excellent" health rating (OR^{good} = 0.87 (0.75, 1.00) and $OR^{fair/poor}$ = 0.83 (0.71, 0.98), AUC = 0.52). BMI was associated with participation (AUC = 0.53). Participants with a BMI greater or equal to 30 kg/m^2 had increased odds of participation compared to those with a BMI less than 25 (OR = 1.22 (1.09, 1.37). Smoking status was also associated with participation (AUC = 0.53). Current smokers had a decreased odds of participation compared to participants that had never smoked (OR = 0.67 (0.59, 0.76)).

In model 1, Black, Chinese, and Japanese women had an increased odds of participation while Hispanic women had a decreased odds of participation compared to White women (range from $OR^{Chinese}$ = 3.51 (2.70, 4.58) to $OR^{Hispanic}$ = 0.84 (0.70, 1.00); AUC = 0.65) with a range of

predictors are included. Model 2 adds an interaction between racial/ethnic group and education (likelihood ratio test: $X^2 = 80.08$ (p-value <0.001), interaction coefficients = 0: $X^2 = 74.99$ (p-value <0.001), AUC = 0.66). White women showed a decrease in odds of participation with decreases in education ($OR^{\text{some college}} = 0.76$ (0.64, 0.91) and $OR^{<=HS} = 0.31$ (0.25, 0.37), reference is >= college degree) while other groups did not exhibit this trend. Model 3 adds an interaction between racial/ethnic group and financial hardship (likelihood ratio test: $X^2 = 31.03$ (p-value <0.001), interaction coefficients = 0: $X^2 = 28.76$ (p-value <0.001), AUC = 0.66). Increased financial hardship was associated with decreased odds of eligibility for White women ($OR^{\text{somewhat hard}} = 0.84$ (0.71, 0.99) and $OR^{\text{very hard}} = 0.77$ (0.58, 1.04), reference was "not very hard") but not for other racial/ethnic groups.

2.4 Discussion

This paper is among the first to document selection mechanisms into a cohort and assess the potential for differential selection and left truncation by racial/ethnic group in a midlife-aging cohort of women. Two mechanisms of selection were identified from the cross-sectional screening study – *eligibility* for the cohort and among those eligible *participation* in the cohort. Eligibility differed by racial/ethnic group, with left truncation mostly stemming from high rates of surgical menopause among Black and Hispanic women. Health related predictors were strongly associated with eligibility while participation was mostly associated with demographic characteristics, thus eligibility may be a stronger source of left truncation than participation. Results point to a potential for bias in estimates of racial/ethnic disparities in reproductive aging and related outcomes in the SWAN cohort and provide a framework for examining selection bias at study commencement in cohorts of aging.

The SWAN study, similar to other cohorts of reproductive and general aging[17,20–25], was subject to two distinct selection mechanisms at study commencement that can be tracked through the cohort screening study. The first was being eligible for the cohort according to the selection criteria, which for the SWAN cohort was based on a woman's reproductive stage. Nested within the eligibility mechanism is the inherent study design that, like many cohorts[17,20–25], selects persons within an eligible age range. Probability of eligibility can vary dependent on the age that a person happens to be when recruited for the study. For example, eligibility in SWAN was

based primarily on where a woman is in her reproductive lifecycle thus odds of eligibility decreased with increasing age. The second selection mechanism, given eligibility, was willingness and ability to participate in the cohort. Many cohort studies present participation rates among eligible participants, providing a summary of how the enrolled cohort differs (or not) from the eligible population and/or engage in complex sampling design to account for differential probability of selection to improve generalizability but ignore eligibility differences[17,20–25]. However, the selected sample is influenced not only by who is approached and willing to participate but also by who is considered eligible.

In this study, rates of eligibility differed by racial/ethnic group. Black women were the least likely to be eligible for the cohort followed by White, Chinese, Hispanic, and then Japanese women. Black women had nearly two times the rate of ineligibility due to surgical menopause (i.e. hysterectomy and/or oophorectomy) compared to other racial/ethnic groups, consistent with other studies showing a higher rate of surgical menopause in Black versus White women in early midlife and midlife[26–29]. Notably the rates of ineligibility due to being in late-peri or naturally post menopause were similar across racial groups, consistent with some but not all previous findings suggesting little difference in the timing of natural menopause across racial groups[18,28,30–37]. Age was an important predictor of eligibility and Black women had a significantly lower odds of eligibility with increasing age compared to White women consistent with the theory that "weathered" age may not be equal between racial groups[13,16]. Education was also an important predictor of eligibility. Among White women decreasing levels of education were associated with lower odds of eligibility, however the trend was not apparent for other racial/ethnic groups. The heterogenous effect may be due to higher levels of education leading to an increase health-care access where minoritized women, who have higher prevalence of reproductive conditions such as fibroids[27,38–40] and may be more likely to have been treated via hysterectomy and/or oophorectomy[27]. Trends were dissimilar for participation, once eligible, Japanese, Hispanic, and Black women were most likely to participate in the cohort followed by White and Chinese women. Similar to other studies[20,41,42], women with higher education and less financial hardship were more likely to participate in the cohort. Given differential associations with racial/ethnic group and other predictors, these results highlight the importance of distinguishing between eligibility and participation mechanisms.

When eligibility criteria are not based on an outcome of interest, left truncation may not be induced. Selection criteria (e.g., where a person lives, # of household members) may not be related to the exposure/outcome relationship. However, in cohorts such as SWAN, where eligibility criteria are based on a health related outcome of interest (i.e. early stages of the menopausal transition) left-truncation is induced by design in the eligibility selection mechanism[1]. As a result, reproductive and health characteristics were significant predictors of eligibility for the SWAN cohort. Selective participation can at least affect the generalizability of a sample to the target and/or general population, and at worst may bias estimates if participation is related to the exposure-outcome association[41–46]. For example, "healthy volunteer" bias can cause those with the outcome to not take part in the study inducing left-truncation[41,42]. However, in SWAN most health characteristics were not significantly associated with participation, showing little support for this type of bias. Cohort studies formed out of national registries[41,42], have found that odds ratios for selected exposure/outcome associations were not largely biased by non-participation[42].

There is a potential for differential selection, particularly from the eligibility selection mechanism to bias estimates of racial/ethnic differences in the SWAN cohort. When estimating racial/ethnic differences in either menopausal timing or other related outcomes in SWAN (Figure 3), sociodemographic, reproductive and health predictors (Z) are a common causes of both selection into the study, and the outcome causing collider stratification bias[3]. This collider stratification bias is like simulation models in work by Mayeda et al. which found that the magnitude of the bias was modest when the common causes (Z) were on the causal pathway between exposure and outcome and is largest when there is an interaction (not on the causal pathway) between exposure and the common causes (Z). The presence of interaction between racial/ethnic group and strong predictors of selection such as education (figure 3, causal structure 2) and age (figure 3, causal structure 2) in the SWAN study would potentially lead to strong biases in the estimation of racial/ethnic differences in the occurrence of final menstrual period and related outcomes, larger than a case where interactions were not present[4,7,47–49]. Cain et al. used simulations to show how increasing amounts of left truncation lead to higher estimates of the age of menopause and how left truncation on menopausal timing can bias estimates on

related outcomes[1]. Taken together, results lend support for the potential to underestimate the racial/ethnic differences in menopause and other conditions influenced by menopausal timing in the SWAN study.

This study has some limitations. We do not have information on how the target population at each site differed from the cross-sectional screening participants. Therefore, generalizability to the target population and/or lack of additional bias stemming from selection into the cross-sectional study cannot be assumed to be negligible. Racial/ethnic groups, by design, are conflated with study site differences in SWAN as three racial/ethnic groups were only included at one study site each.

In conclusion, results show two distinct selection mechanisms and that for the SWAN study "healthy volunteer" bias captured in participation rates may be less important in non-occupational cohorts[41], while eligibility criteria may influence[22] selection bias and left-truncation more. Eligibility into the SWAN cohort differed by racial/ethnic group, with Black women least likely to be eligible for the cohort due to nearly a third of Black women with surgical menopause. Combined with interactions between racial/ethnic group, sociodemographic and health in predicting eligibility into the cohort, results present a strong case that current estimates that do not account for selection at study commencement and left truncation may underestimate racial/ethnic disparities in menopausal timing and other related outcomes. Ability to precisely estimate the probability of left truncation using 1) the cross-sectional recruitment study and 2) from within the cohort, the time to pre-peri menopause and/or surgical menopause, is a unique strength of the SWAN cohort. Further research will pursue using this information to account for all forms of truncation and selection bias. This analysis provides a framework and important first steps for SWAN and other cohorts of aging to examine and quantify the potential bias caused by selection at study commencement.

Figure 2.1 Selection Mechanisms from the Cross-Sectional Screening Study

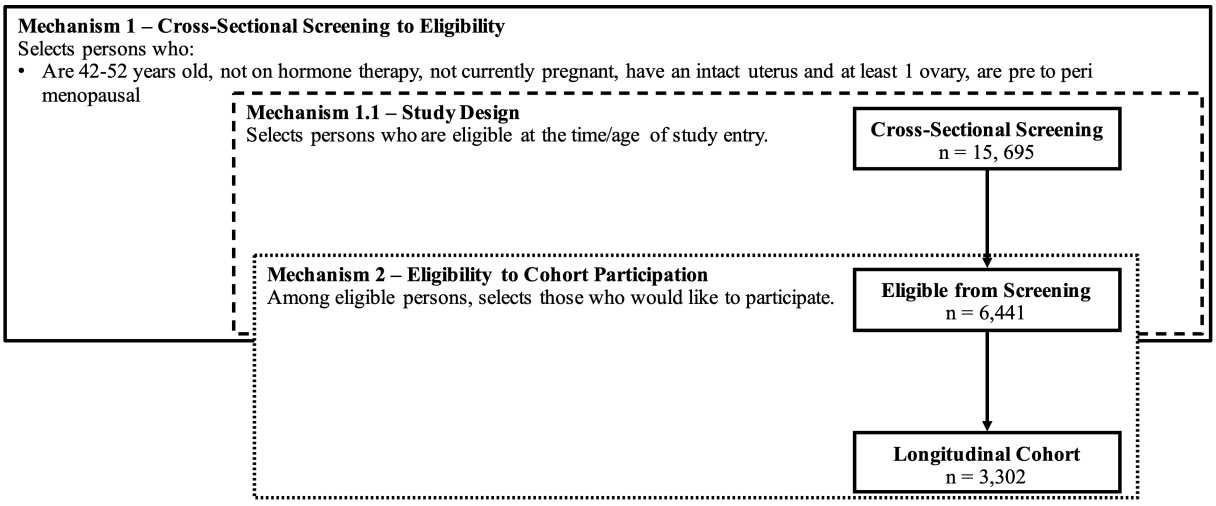


Table 2.1 Demographics at Each Stage of Selection from the Cross-Sectional Screening Study

	Cross-Sectional Screening n = 15,695		**Eligible Participants** n = 6,521		**Cohort at Baseline** n = 3,302	
Demographic Characteristics						
Racial/Ethnic Group						
Black	4,402	28.1	1,709	26.2	935	28.3
White	7,805	49.7	3,204	49.1	1,551	47.0
Chinese	653	4.2	362	5.6	250	7.6
Hispanic	1,979	12.6	802	12.3	286	8.7
Japanese	856	5.5	444	6.8	281	8.5
Site						
Michigan	2,563	16.3	922	14.1	543	16.4
Boston	2,137	13.6	909	13.9	452	13.7
Chicago	1,406	9.0	616	9.5	457	13.8
Davis	1,484	9.5	685	10.5	459	13.9
Los Angeles	2,200	14.0	932	14.3	496	15.0
New Jersey	3,307	21.1	1,383	21.2	432	13.1
Pittsburgh	2,598	16.6	1,074	16.5	464	14.1
Age	15,694	47.6 (4.3)	6,520	46.2 (2.7)	3,302	46.2 (2.7)
Missing	*1*	*0.0*	*1*	*0.0*		-
Educational Level						
<= High School	5,637	35.9	2,120	32.5	825	25.0
Some College	4,935	31.4	2,015	30.9	1,070	32.4
>= College	5,078	32.4	2,358	36.2	1,408	42.6
Missing	*45*	*0.3*	*28*	*0.4*		-
Marital Status						
Never	1,982	12.6	881	13.5	446	13.5
Previously	3,799	24.2	1,387	21.3	669	20.3
Currently	9,885	63.0	4,244	65.1	2,188	66.2
Missing	*29*	*0.2*	*9*	*0.1*		-
Financial Hardship						
Very Hard	1,825	11.6	706	10.8	306	9.3
Somewhat Hard	5,091	32.4	2,129	32.7	1,014	30.7
Not Very Hard	8,725	55.6	3,670	56.3	1,983	60.0
Missing	*54*	*0.3*	*16*	*0.3*		-
Reproductive Characteristics						
Ever Used Hormonal Birth Control	12,124	77.3	4,759	73.0	2,521	76.3
Missing	*18*	*0.1*	*1*	*0.0*		-
Parity						
0	2,484	15.8	1,044	16.0	541	16.4
1	2,629	16.8	1,102	16.9	538	16.3
2	5,099	32.5	2,136	32.8	1,079	32.7
3	3,095	19.7	1,300	19.9	625	18.9
4 +	2,388	15.2	939	14.4	520	15.7
Missing	*0*	*0.0*	*0*	*0.0*		-
Ever Diagnosed with Fibroids	4,013	25.6	1,232	18.9	674	20.4
Missing	*51*	*0.3*	*18*	*0.3*		-
Overall Health Characteristics						
Self-Rated Health						
Excellent	2,974	19.0	1,319	20.2	689	20.9
Very Good	4,972	31.7	2,214	34.0	1,171	35.5
Good	4,791	30.5	1,927	29.6	937	28.4
Fair/Poor	2,958	18.9	1,061	16.3	506	15.3

Missing	*0*	*0.0*	*0*	*0.0*		-
Ever Diagnosed with Diabetes	1,028	6.6	329	5.1	163	4.9
Missing	*9*	*0.1*	*2*	*0.0*		-
Ever Diagnosed with Heart Attack/Angina	456	2.9	100	1.5	58	1.8
Missing	*6*	*0.0*	*3*	*0.1*		-
Ever Diagnosed with Osteoporosis	407	2.6	86	1.3	42	1.3
Missing	*31*	*0.2*	*5*	*0.1*		-
Ever Diagnosed with Cancer	655	4.2	122	1.9	62	1.9
Missing	*13*	*0.1*	*3*	*0.1*		-
Body Mass Index						
<25	6,765	43.1	2,936	45.0	1,454	44.0
25-29.9	4,319	27.5	1,693	26.0	817	24.7
>=30	4,611	29.4	1,892	29.0	1,032	31.2
Missing	*0*	*0.0*	*0*	*0.0*		-
Smoking Status						
Never	8,166	52.0	3,626	55.6	1,918	58.1
Former	3,897	24.8	1,545	23.7	816	24.7
Current	3,565	22.7	1,329	20.4	569	17.2
Missing	*67*	*0.4*	*21*	*0.3*		-

Note: No missing among cohort sample due to imputation.

Table 2.2 Eligibility and Participation Proportions by Racial/Ethnic Group from the Cross-Sectional Screening Study

	Total n = 15,695	**Black** n = 4,402	**White** n = 7,805	**Japanese** n = 653	**Hispanic** n = 1,979	**Chinese** n = 856
Eligible %(n)	41.6 (6521)	38.8 (1709)	41.1 (3204)	55.4 (362)	40.5 (802)	51.9 (444)
Ineligible %(n)	58.5 (9174)	61.2 (2693)	60.0 (4601)	44.6 (291)	59.5 (1177)	48.1 (412)
Reasons for Ineligibility (%)[a]						
Currently Pregnant	0.2	0.1	0.3	0.2	0.2	0.2
Surgical Menopause	20.1	30.9	17.0	5.7	17.3	9.7
Late Peri[c] or Natural Post Menopause	4.8	4.2	5.3	4.0	4.6	4.4
Current Hormone Use	18.2	13.8	23.8	14.7	8.9	14.7
Outside Age Range	30.8	29.3	31.8	25.0	32.0	31.8
Participated %(n)	50.7 (3302)	54.7 (934)	48.4 (1551)	69.1 (250)	35.7 (286)	63.3 (281)

Note: [a]Persons can be ineligible for not menstruating (pregnancy, surgical menopause, and natural menopause) and/or current hormone use and/or outside age range, percentages do not add up to 100.
[b]Includes Hysterectomy and Bi-Lateral Oophorectomy.
[c]No period in the past 3 months.

Table 2.3 Bivariate and Multivariate Logistic Regression of Predictors and Eligibility from the Cross-Sectional Screening Study (n = 15,454)

	Individual Bivariate Models -					Model 1 *Multivariate with No Interactions* auc = 0.70				Model 2[a] *Age Interaction* auc = 0.70				Model 3[b] *Educational Level Interaction* auc = 0.70			
	OR	LCI	UCI	p-value	AUC	OR	LCI	UCI	p-value	OR	LCI	UCI	p-value	OR	LCI	UCI	p-value
Demographic Characteristics																	
Racial/Ethnic Group (ref: White)					0.53												
Black	0.91	0.84	0.98	0.016		1.03	0.94	1.12	0.582	3.59	1.40	9.19	0.008	1.02	0.87	1.19	0.828
Chinese	1.79	1.52	2.10	0.000		1.63	1.36	1.96	0.000	0.06	0.00	0.47	0.008	1.42	1.09	1.85	0.010
Hispanic	0.98	0.88	1.08	0.671		0.91	0.81	1.02	0.107	0.42	0.13	1.36	0.146	0.63	0.49	0.82	0.001
Japanese	1.55	1.34	1.78	0.000		1.43	1.22	1.67	0.000	0.23	0.04	1.23	0.085	1.06	0.84	1.33	0.625
Site (ref: Michigan)					0.53												
Boston	1.23	1.09	1.38	0.001													
Chicago	1.40	1.23	1.60	0.000													
Davis	1.49	1.31	1.69	0.000													
Los Angeles	1.28	1.14	1.44	0.000													
New Jersey	1.16	1.05	1.29	0.004													
Pittsburgh	1.26	1.13	1.41	0.000													
Age	0.88	0.88	0.89	0.000	0.64	0.89	0.88	0.89	0.000					0.88	0.88	0.89	0.000
Racial/Ethnic Group x Age					0.65[c]												
White	0.88	0.87	0.89	*ref*						0.89	0.88	0.90	ref				
Black	0.86	0.83	0.89	0.011						0.86	0.84	0.89	0.009				
Chinese	0.96	0.91	1.01	0.000						0.95	0.90	1.01	0.002				
Hispanic	0.90	0.86	0.93	0.179						0.90	0.87	0.93	0.194				
Japanese	0.90	0.86	0.95	0.144						0.92	0.88	0.96	0.033				
Educational Level (ref: >= College)					0.54												
Some College	0.80	0.74	0.86	0.000		0.82	0.75	0.89	0.000	0.81	0.75	0.89	0.000				
<= High School	0.69	0.64	0.74	0.000		0.74	0.68	0.81	0.000	0.74	0.68	0.81	0.000				
Racial/Ethnic Group x Educational Level (ref: >= College)					0.56[d]												
White																	
Some College	0.77	0.69	0.86	*ref*										0.78	0.69	0.87	ref
<= High School	0.61	0.55	0.69											0.65	0.58	0.74	
Black																	
Some College	0.84	0.62	1.13	0.220										0.79	0.57	1.08	0.884
<= High School	0.69	0.51	0.94	0.383										0.68	0.49	0.94	0.757
Chinese																	
Some College	0.85	0.50	1.43	0.019										0.85	0.48	1.51	0.674
<= High School	0.96	0.59	1.55	0.639										0.97	0.57	1.65	0.062
Hispanic																	
Some College	0.94	0.61	1.45	0.000										0.96	0.61	1.52	0.212
<= High School	1.06	0.72	1.57	0.219										1.15	0.76	1.76	0.000
Japanese																	
Some College	1.09	0.71	1.67	0.008										1.24	0.79	1.97	0.007
<= High School	1.02	0.63	1.67	0.033										1.23	0.73	2.09	0.002
Financial Harship (ref = Not Very Hard)					0.51												
Somewhat Hard	0.97	0.91	1.04	0.467													
Very Hard	0.84	0.76	0.93	0.001													
Financial Hardship (ref: Not Very Hard)					0.54[e]												
White																	
Somewhat Hard	0.96	0.87	1.07	*ref*													
Very Hard	0.83	0.70	0.99														
Black																	
Somewhat Hard	1.07	0.82	1.39	0.241													
Very Hard	1.05	0.68	1.61	0.074													
Chinese																	
Somewhat Hard	0.85	0.53	1.36	0.507													
Very Hard	0.66	0.28	1.59	0.528													
Hispanic																	
Somewhat Hard	1.17	0.82	1.66	0.135													
Very Hard	0.97	0.60	1.57	0.299													
Japanese																	
Somewhat Hard	1.16	0.75	1.77	0.279													
Very Hard	1.21	0.46	3.19	0.348													
Marital Status (ref: Currently)					0.53												
Previously	0.76	0.71	0.83	0.000		0.92	0.85	1.01	0.071	0.92	0.85	1.01	0.069	0.93	0.85	1.01	0.088
Never	1.05	0.95	1.15	0.341		1.07	0.95	1.20	0.250	1.05	0.94	1.18	0.384	1.07	0.96	1.20	0.234
Reproductive Characteristics																	
Ever Used Hormonal Birth Control	0.68	0.63	0.74	0.000	0.53	0.66	0.60	0.71	0.000	0.65	0.60	0.71	0.000	0.66	0.61	0.72	0.000
Parity (ref: 0)					0.51												
1	1.01	0.90	1.12	0.905		1.20	1.06	1.36	0.004	1.19	1.05	1.35	0.006	1.21	1.07	1.37	0.003
2	1.01	0.92	1.11	0.843		1.17	1.05	1.31	0.006	1.16	1.03	1.30	0.011	1.18	1.05	1.32	0.004
3	1.01	0.91	1.12	0.853		1.25	1.11	1.42	0.000	1.25	1.10	1.41	0.001	1.26	1.12	1.43	0.000
4 +	0.90	0.80	1.00	0.058		1.30	1.14	1.49	0.000	1.30	1.13	1.48	0.000	1.31	1.14	1.50	0.000
Ever Diagnosed with Fibroids	0.54	0.50	0.58	0.000	0.56	0.59	0.55	0.64	0.000	0.59	0.54	0.64	0.000	0.59	0.54	0.64	0.000
Overall Health Characteristics																	
Self-Rated Health (ref: Excellent)					0.54												
Very Good	1.01	0.92	1.11	0.813		1.04	0.94	1.14	0.485	1.04	0.94	1.15	0.447	1.04	0.94	1.15	0.439
Good	0.84	0.77	0.92	0.000		0.90	0.81	1.00	0.040	0.90	0.81	1.00	0.050	0.90	0.81	1.00	0.051
Fair/Poor	0.70	0.63	0.78	0.000		0.85	0.75	0.96	0.011	0.86	0.76	0.97	0.013	0.84	0.75	0.95	0.007
Ever Diagnosed with Diabetes	0.65	0.56	0.74	0.000	0.51												
Ever Diagnosed with Heart Attack/Angina	0.39	0.31	0.49	0.000	0.51	0.58	0.46	0.74	0.000	0.59	0.46	0.75	0.000	0.59	0.47	0.76	0.000
Ever Diagnosed with Osteoporosis	0.37	0.29	0.47	0.000	0.51	0.57	0.44	0.74	0.000	0.57	0.44	0.73	0.000	0.57	0.44	0.74	0.000
Ever Diagnosed with Cancer	0.31	0.25	0.38	0.000	0.52	0.37	0.30	0.46	0.000	0.37	0.30	0.46	0.000	0.37	0.30	0.46	0.000
Body Mass Index (ref: <25)					0.52												
25-29.9	0.83	0.77	0.90	0.000		0.98	0.90	1.07	0.618	0.98	0.90	1.07	0.627	0.98	0.90	1.07	0.673
>=30	0.91	0.84	0.98	0.015		1.16	1.06	1.27	0.001	1.17	1.07	1.27	0.001	1.17	1.07	1.28	0.000
Smoking Status (ref: Never)					0.53												
Former	0.84	0.78	0.91	0.000		0.97	0.89	1.06	0.478	0.97	0.89	1.06	0.503	0.97	0.89	1.05	0.432
Current	0.76	0.70	0.82	0.000		0.84	0.77	0.92	0.000	0.84	0.77	0.92	0.000	0.85	0.78	0.94	0.001

Note: [a]Model 2 (Interaction) versus Model 1: X2 = 79.047 (< 0.001). Racial/Ethnic group x age interaction = 0: X2 = 27.90 (< 0.001).
[b]Model 3 (Interaction) versus Model 1: X2 = 76.287 (< 0.001). Racial/Ethnic group x age interaction = 0: X2 = 29.43 (< 0.001).
[c]Bivariate Racial/Ethnic group x age interaction = 0: X2 = 31.39 (< 0.001).
[d]Bivariate Racial/Ethnic group x educational level interaction = 0: X2 = 27.51 (< 0.001).
[e]Bivariate Racial/Ethnic group x financial hardship interaction = 0: X2 = 8.13 (0.4209).

Figure 2.2 Predicted Probability of Eligibility by Age at Study Screening across Racial/Ethnic Groups in the Cross-Sectional Screening Study

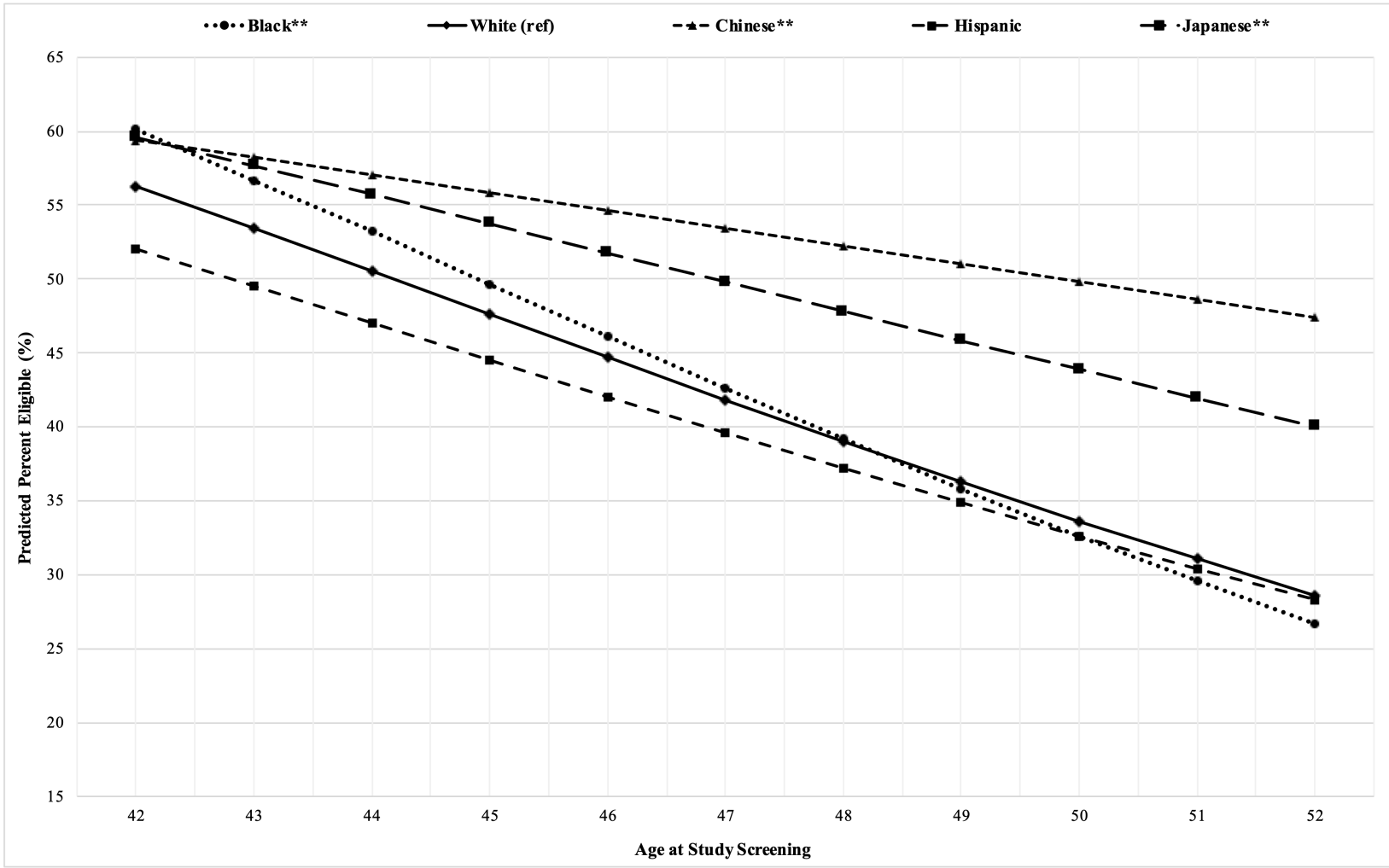
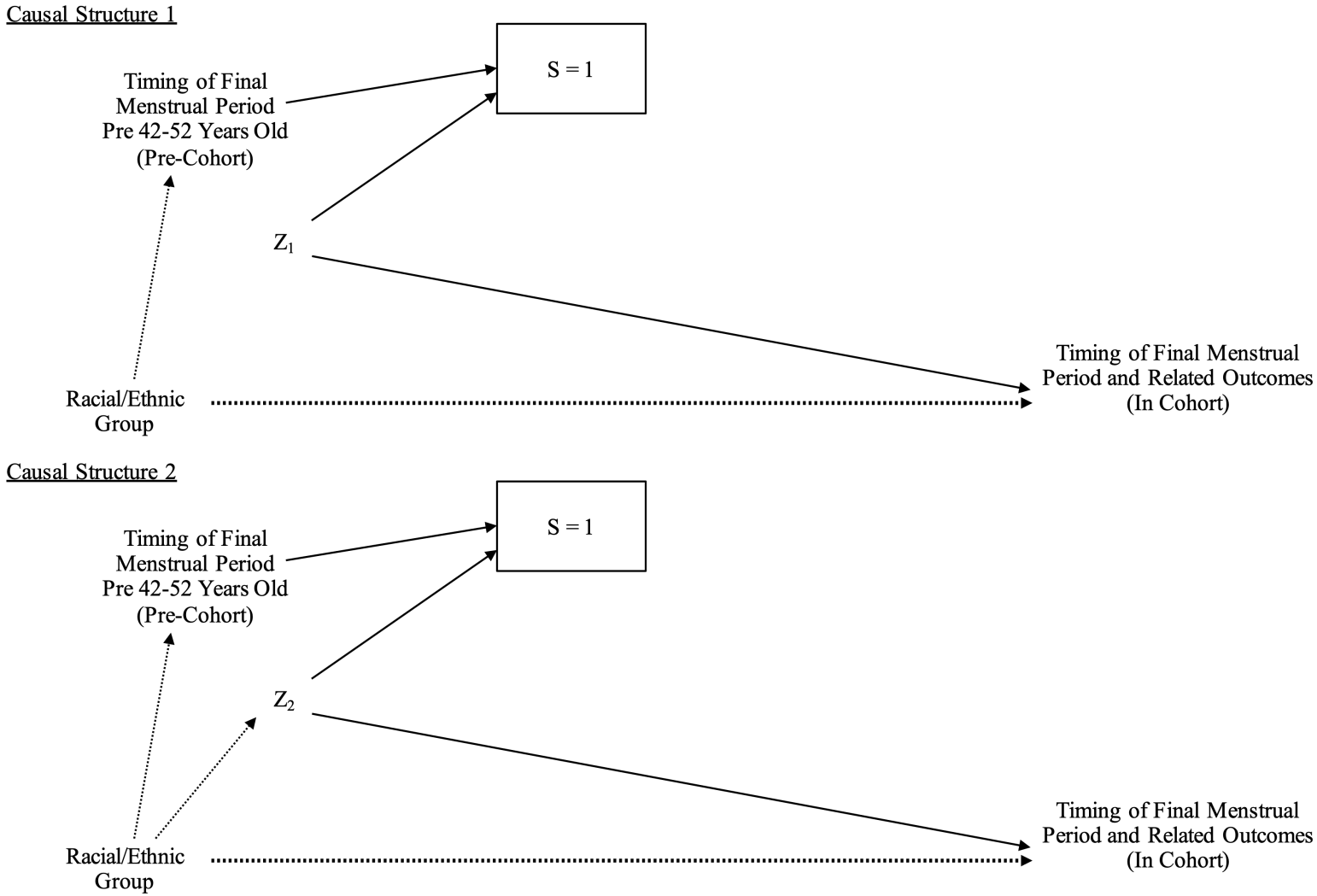


Table 2.4 Bivariate and Multivariate Logistic Regression of Predictors and Participation among Eligible Women from the Cross-Sectional Screening Study (n = 6,521)

	Individual Bivariate Models -					Model 1 *Multivariate with No Interactions* auc = 0.65				Model 2[a] *Educational Level Interaction* auc = 0.66				Model 3[b] *Financial Hardship Interaction* auc = 0.66			
	OR	LCI	UCI	p-value	AUC	OR	LCI	UCI	p-value	OR	LCI	UCI	p-value	OR	LCI	UCI	p-value
Demographic Characteristics																	
Racial/Ethnic Group (ref: White)					0.58												
Black	1.29	1.14	1.45	0.000		1.36	1.19	1.55	0.000	1.28	1.02	1.61	0.033	1.38	1.16	1.65	0.000
Chinese	2.38	1.89	3.01	0.000		3.51	2.70	4.58	0.000	2.87	1.94	4.25	0.000	3.14	2.30	4.28	0.000
Hispanic	0.59	0.50	0.69	0.000		0.84	0.70	1.00	0.055	0.24	0.15	0.38	0.000	0.46	0.32	0.65	0.000
Japanese	1.84	1.50	2.26	0.000		2.23	1.79	2.78	0.000	1.97	1.41	2.74	0.000	2.04	1.57	2.65	0.000
Site (ref: Michigan)					0.65												
Boston	0.69	0.57	0.83	0.000													
Chicago	2.01	1.61	2.51	0.000													
Davis	1.42	1.15	1.74	0.001													
Los Angeles	0.79	0.66	0.95	0.014													
New Jersey	0.32	0.27	0.38	0.000													
Pittsburgh	0.53	0.44	0.63	0.000													
Age	0.99	0.98	1.01	0.569	0.50	0.99	0.97	1.01	0.207	0.98	0.97	1.00	0.114				
Racial/Ethnic Group x Age					0.58[c]												
White	1.00	0.97	1.02	*ref*													
Black	0.97	0.91	1.04	0.173													
Chinese	1.01	0.90	1.13	0.815													
Hispanic	1.01	0.93	1.10	0.761													
Japanese	0.97	0.88	1.08	0.514													
Educational Level (ref: >= College)					0.59									0.99	0.97	1.01	0.195
Some College	0.76	0.68	0.86	0.000		0.75	0.66	0.86	0.000					0.76	0.67	0.86	0.000
<= High School	0.43	0.38	0.48	0.000		0.45	0.40	0.52	0.000					0.45	0.39	0.52	0.000
Racial/Ethnic Group x Educational Level					0.64[d]												
White																	
Some College	0.78	0.66	0.92	*ref*						0.76	0.64	0.91	*ref*				
<= High School	0.31	0.26	0.37							0.31	0.25	0.37					
Black																	
Some College	0.71	0.45	1.12	0.538						0.71	0.44	1.14	0.629				
<= High School	0.43	0.26	0.71	0.035						0.44	0.26	0.73	0.030				
Chinese																	
Some College	0.69	0.31	1.52	0.711						0.83	0.36	1.92	0.797				
<= High School	0.42	0.21	0.88	0.259						0.59	0.28	1.27	0.024				
Hispanic																	
Some College	1.75	0.84	3.63	0.005						1.63	0.77	3.43	0.010				
<= High School	2.25	1.13	4.47	0.000						2.11	1.04	4.27	0.000				
Japanese																	
Some College	0.61	0.32	1.15	0.317						0.66	0.35	1.26	0.555				
<= High School	0.61	0.29	1.28	0.016						0.72	0.34	1.53	0.003				
Financial Harship (ref = Not Very Hard)					0.54												
Somewhat Hard	0.77	0.70	0.86	0.000		0.92	0.81	1.03	0.147	0.91	0.81	1.03	0.131				
Very Hard	0.65	0.55	0.77	0.000		0.92	0.76	1.11	0.386	0.91	0.75	1.09	0.305				
Racial/Ethnic Group x Financial Hardship (ref = Not Very Hard)					0.60[e]												
White																	
Somewhat Hard	0.78	0.67	0.91	*ref*										0.84	0.71	0.99	*ref*
Very Hard	0.64	0.48	0.84											0.77	0.58	1.04	
Black																	
Somewhat Hard	0.73	0.48	1.11	0.638										0.77	0.50	1.18	0.525
Very Hard	0.82	0.41	1.62	0.234										0.96	0.47	1.97	0.304
Chinese																	
Somewhat Hard	0.68	0.34	1.35	0.617										1.00	0.48	2.08	0.540
Very Hard	1.32	0.30	5.70	0.230										4.42	0.70	28.03	0.028
Hispanic																	
Somewhat Hard	1.88	1.06	3.34	0.000										2.16	1.20	3.89	0.000
Very Hard	1.32	0.60	2.91	0.005										1.56	0.69	3.53	0.009
Japanese																	
Somewhat Hard	0.97	0.52	1.82	0.345										1.10	0.58	2.09	0.255
Very Hard	1.14	0.28	4.69	0.313										1.13	0.26	4.81	0.522
Marital Status (ref: Currently)					0.51												
Previously	0.88	0.78	0.99	0.032		0.92	0.80	1.05	0.206	0.93	0.81	1.06	0.286	0.93	0.81	1.06	0.269
Never	0.96	0.83	1.11	0.615		0.96	0.81	1.14	0.648	0.96	0.81	1.14	0.630	0.98	0.82	1.16	0.788
Reproductive Characteristics																	
Ever Used Hormonal Birth Control	1.41	1.26	1.57	0.000	0.53	1.38	1.22	1.56	0.000	1.39	1.23	1.58	0.000	1.39	1.23	1.57	0.000
Parity (ref: 0)					0.52												
1	0.89	0.75	1.05	0.165		0.92	0.77	1.11	0.402	0.93	0.77	1.13	0.472	0.93	0.77	1.12	0.455
2	0.95	0.82	1.10	0.489		1.00	0.84	1.18	0.998	1.01	0.85	1.20	0.881	1.01	0.85	1.20	0.919
3	0.86	0.73	1.01	0.072		0.99	0.83	1.20	0.953	1.00	0.83	1.21	0.988	0.99	0.82	1.20	0.949
4 +	1.15	0.97	1.38	0.113		1.41	1.15	1.74	0.001	1.39	1.13	1.70	0.002	1.43	1.16	1.75	0.001
Ever Diagnosed with Fibroids	1.21	1.07	1.38	0.002	0.51												
Overall Health Characteristics																	
Self-Rated Health (ref: Excellent)					0.52												
Very Good	1.03	0.90	1.18	0.706		1.05	0.91	1.21	0.505	1.05	0.91	1.22	0.470	1.05	0.91	1.22	0.471
Good	0.87	0.75	1.00	0.043		0.94	0.80	1.09	0.407	0.95	0.81	1.11	0.531	0.94	0.80	1.10	0.434
Fair/Poor	0.83	0.71	0.98	0.028		0.94	0.78	1.13	0.514	0.92	0.76	1.11	0.377	0.92	0.77	1.12	0.412
Ever Diagnosed with Diabetes	0.95	0.76	1.19	0.676	0.50												
Ever Diagnosed with Heart Attack/Angina	1.35	0.91	2.03	0.141	0.50												
Ever Diagnosed with Osteoporosis	0.93	0.60	1.42	0.729	0.50												
Ever Diagnosed with Cancer	1.01	0.70	1.44	0.974	0.50												
Body Mass Index (ref: <25)					0.53												
25-29.9	0.95	0.84	1.07	0.407		1.09	0.96	1.25	0.172	1.11	0.97	1.26	0.131	1.10	0.96	1.25	0.161
>=30	1.22	1.09	1.37	0.001		1.59	1.39	1.82	0.000	1.62	1.41	1.85	0.000	1.60	1.40	1.83	0.000
Smoking Status (ref: Never)					0.53												
Former	1.00	0.88	1.12	0.958		1.05	0.93	1.20	0.417	1.05	0.92	1.19	0.477	1.06	0.93	1.20	0.393
Current	0.67	0.59	0.76	0.000		0.81	0.71	0.93	0.003	0.84	0.73	0.97	0.016	0.82	0.71	0.94	0.004

Note: [a]Model 2 (Interaction) versus Model 1: X2 = 80.48 (< 0.001). Racial/Ethnic group x educational level interaction = 0: X2 = 74.99 (< 0.001).
[b]Model 3 (Interaction) versus Model 1: X2 = 31.03 (< 0.001). Racial/Ethnic group x financial hardship interaction = 0: X2 = 28.76 (< 0.001).
[c]Bivariate Racial/Ethnic group x age interaction = 0: X2 = 2.70 (0.609).
[d]Bivariate Racial/Ethnic group x educational level interaction = 0: X2 = 77.64 (< 0.001).
[e]Bivariate Racial/Ethnic group x financial hardship interaction = 0: X2 = 24.18 (<0.005).

Figure 2.3 Proposed Causal Structures of the Effect of Selection at Study Commencement on the Association between Racial/Ethnic Group and Main Outcomes



Note: Timing of final menstrual period includes surgical (hysterectomy/oophorectomy) and natural menopause. Z_1 represents a factor or set of factors that influence the main outcomes of the study as well as selection and are not on the causal pathway between

racial/ethnic group and the main outcomes, examples could include chronologic age, parity, birth control use. Z_2 represents a factor or set of factors that influence the main outcomes of the study as well as selection and are on the causal pathway between racial/ethnic group and the main outcomes. Examples could include socioeconomic status and chronic health conditions.

2.5 References

1. Cain, K. C. *et al.* Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes. *Am. J. Epidemiol.* **173**, 1078–1084 (2011).
2. Schisterman, E. F., Cole, S. R., Ye, A. & Platt, R. W. Accuracy loss due to selection bias in cohort studies with left truncation. *Paediatr. Perinat. Epidemiol.* **27**, 491–502 (2013).
3. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A structural approach to selection bias. *Epidemiology* **15**, 615–625 (2004).
4. Mayeda, E. R. *et al.* A Simulation Platform for Quantifying Survival Bias: An Application to Research on Determinants of Cognitive Decline. *Am. J. Epidemiol.* **184**, 378–387 (2016).
5. Howe, C. J., Cole, S. R., Lau, B., Napravnik, S. & Eron, J. J. Selection bias due to loss to follow up in cohort studies. *Epidemiology* **176**, 139–148 (2017).
6. Howards, P. P., Hertz-Picciotto, I. & Poole, C. Conditions for bias from differential left truncation. *Am. J. Epidemiol.* **165**, 444–452 (2007).
7. Mayeda, E. R., Filshtein, T. J., Tripodis, Y., Maria Glymour, M. & Gross, A. L. Does selective survival before study enrolment attenuate estimated effects of education on rate of cognitive decline in older adults? A simulation approach for quantifying survival bias in life course epidemiology. *Int. J. Epidemiol.* **47**, 1507–1517 (2018).
8. Kochanek, K. D., Arias, E. & Anderson, R. N. *How Did Cause of Death Contribute to Racial Differences in Life Expectancy in the United States in 2010? NCHS data brief, no 125* (2013).
9. Nan, |, Astone, M., Ensminger, M. & Juon, H. S. *Early Adult Characteristics and Mortality Among Inner-City African-American Women*. *American Journal of Public Health* **92**, (2002).
10. Geronimus, A. T., Bound, J., Waidmann, T. A., Hillemeier, M. M. & Burns, P. B. Excess Mortality among Blacks and Whites in the United States. *N. Engl. J. Med.* **335**, 1552–1558 (1996).
11. Geronimus, A. T., Bound, J., Waidmann, T. A., Colen, C. G. & Steffick, D. Inequality in Life Expectancy, Funcational Status, and Active Life Expectancy Across Selected Black and White Populations in the United States. *Demography* **38**, 227–251 (2001).
12. Elo, I. T. & Preston, S. H. *Educational Differential in Mortality: United States*. *Soc. Sci. Med* **42**, (1996).
13. Geronimus, A. T., Hicken, M., Keene, D. & Bound, J. 'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am. J. Public Health* **96**, 826–833 (2006).
14. Levine, M. E. & Crimmins, E. M. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc. Sci. Med.* **118**, 27–32 (2014).
15. Geronimus, A. T. *et al.* Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-based Sample. *J. Health Soc. Behav.* **56**, 199–224 (2015).
16. Gee, G. C., Hing, A., Mohammed, S., Tabor, D. C. & Williams, D. R. Racism and the Life Course: Taking Time Seriously. *Am J Public Heal.* **109**, 43–47 (2019).
17. Sowers, M. *et al.* SWAN: A Multicenter, Multiethnic, Community Based Cohort Study of Women and the Menopausal Transition. in *Menopause: Biology and Pathobiology* 175–188 (2000).

18. Gold, E. B. *et al.* Factors Associated with Age at Natural Menopause in a Multiethnic Sample of Midlife Women. *Am. J. Epidemiol.* **153**, 865–874 (2001).
19. Rubin, D. B. Multiple Imputations in Sample Surveys - A Phenomenological Baysian Approach to Nonresponse. *Proc. Surv. Res. methods Sect. Am. Stat. Assoc. Am. Stat. Assoc.* **1**, 20–34 (1978).
20. Friedman, G. D. *et al.* Cardia: study design, recruitment, and some characteristics of the examined subjects. *J. Clin. Epidemiol.* **41**, 1105–1116 (1988).
21. Sonnega, A. *et al.* Cohort profile: The Health and Retirement Study (HRS). *Int. J. Epidemiol.* **43**, 576–585 (2014).
22. Ferrell, R. J. & Sowers, M. Longitudinal, epidemiologic studies of female reproductive aging. *Ann. N. Y. Acad. Sci.* **1204**, 188–197 (2010).
23. Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: Objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
24. Hill, C. *et al.* The Atherosclerosis Risk in Communities (ARIC) Study: Design and Objectives. **129**, (1989).
25. Howard, V. J. *et al.* The reasons for geographic and racial differences in stroke study: Objectives and design. *Neuroepidemiology* **25**, 135–143 (2005).
26. Bower, J. K., Schreiner, P. J., Sternfeld, B. & Lewis, C. E. Black-white differences in hysterectomy prevalence: The CARDIA study. *Am. J. Public Health* **99**, 300–307 (2009).
27. Powell, L. H. *et al.* Ethnic differences in past hysterectomy for benign conditions. *Women's Heal. Issues* **15**, 179–186 (2005).
28. McKnight, K. K. *et al.* Racial and regional differences in age at menopause in the United States: findings from the REasons for Geographic And Racial Differences in Stroke (REGARDS) study. *Am. J. Obstet. Gynecol.* **205**, 353.e1-353.e8 (2011).
29. WRIGHT, M. A. *et al.* Changing trends in Black-White racial differences in surgical menopause: a population-based study. *Am. J. Obstet. Gynecol.* (2021). doi:10.1016/j.ajog.2021.05.045
30. Appiah, D., Nwabuo, C. C., Ebong, I. A., Wellons, M. F. & Winters, S. J. Trends in Age at Natural Menopause and Reproductive Life SpanAmongUSWomen, 1959-2018. *JAMA - J. Am. Med. Assoc.* **325**, 1328–1330 (2021).
31. MacMahon, B. & Worcester, J. *Age at menopause. United States--1960-1962. Vital and health statistics. Series 11, Data from the National Health Survey* (1966).
32. Luborsky, J. L., Meyer, P., Sowers, M. F., Gold, E. B. & Santoro, N. Premature menopause in a multi-ethnic population study of the menopause transition*. *Hum. Reprod.* **18**, 199–206 (2002).
33. Bromberger, J. T. *et al.* Prospective Study of the Determinants of Age at Menopause. *Am. J. Epidemiol. Copyr. O* **145**, (1997).
34. Stanford, J. L., Hartge, P., Brinton, L. A., Hoover', R. N. & Brookmeyer, R. Factors Influencing the Age of Natural Menopause. *J. Chronic Dis.* **40**, 995–1002 (1987).
35. Butts, S. F. & Seifer, D. B. Racial and ethnic differences in reproductive potential across the life cycle. *Fertil. Steril.* **93**, 681–690 (2010).
36. Delellis Henderson, K., Bernstein, L., Henderson, B., Kolonel, L. & Pike, M. C. Original Contribution Predictors of the Timing of Natural Menopause in the Multiethnic Cohort Study. *Am. J. Epidemiol.* **167**, 1287–1294 (2008).
37. Gold, E. B. *et al.* Factors related to age at natural menopause: Longitudinal analyses from SWAN. *Am. J. Epidemiol.* (2013). doi:10.1093/aje/kws421

38. Yu, O. *et al.* A US Population-based Study of Uterine Fibroid Diagnosis Incidence, Trends, and Prevalence: 2005-2014. *Am. J. Obstet. Gynecol.* (2018). doi:10.1016/J.AJOG.2018.09.039
39. Stewart, E. A., Nicholson, W. K., Bradley, L. & Borah, B. J. The burden of uterine fibroids for african-american women: Results of a national survey. *J. Women's Heal.* **22**, 807–816 (2013).
40. Marsh, E. E. *et al.* Racial differences in fibroid prevalence and ultrasound findings in asymptomatic young women (18-30 years old): a pilot study. *Fertil. Steril.* **99**, 1951–1957 (2013).
41. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
42. Nohr, E. A., Frydenberg, M., Henriksen, T. B. & Olsen, J. Does low participation in cohort studies induce bias? *Epidemiology* **17**, 413–418 (2006).
43. Rothman, K. J., Gallacher, J. E. J. & Hatch, E. E. Why representativeness should be avoided. *Int. J. Epidemiol.* **42**, 1012–1014 (2013).
44. Richiardi, L., Pizzi, C. & Pearce, N. Commentary: Representativeness is usually not necessary and often should be avoided. *Int. J. Epidemiol.* **42**, 1018–1022 (2013).
45. Richiardi, L., Pizzi, C. & Pearce, N. Representativeness. *Int. J. Epidemiol.* **43**, 632–633 (2014).
46. Stang, A. & Jöckel, K. H. Avoidance of representativeness in presence of effect modification. *Int. J. Epidemiol.* **43**, 630–631 (2014).
47. Greenland, S. Response and Follow-up Bias in Cohort Studies. *Am. J. Epidemiol.* **185**, 1044–1047 (2017).
48. Hernán, M. A. Invited commentary: Selection bias without colliders. *Am. J. Epidemiol.* **185**, 1048–1050 (2017).
49. Glymour, M. M. & Vittinghoff, E. Selection bias as an explanation for the obesity paradox: Just because it's possible doesn't mean it's plausible. *Epidemiology* **25**, 4–6 (2014).

Chapter 3 Racial Differences in Reproductive Aging – The Effect of Multiple Sources of Selection Bias in a Multi-Ethnic Cohort of Aging

3.1 Introduction

Pervasive and consistent evidence of earlier morbidity, mortality and lower life expectancy among racial/ethnic minorities in the US, particularly Blacks compared to Whites, underlines the potential for accelerated aging or "weathering" in these populations[1–7]. "Weathering" defined as early health deterioration due to cumulative impact of repeated experience with social or economic adversity and political marginalization, is a potential fundamental cause of racial/ethnic disparities in morbidity and mortality[7–9]. The consequence of "weathering" are seen in the high rates of cardio-metabolic diseases such as heart disease, stroke and diabetes as leading causes of racial/ethnic disparities in lifespan[3]. Studies have found that various cardio-metabolic risk factors that are indicators of aging start to accumulate earlier in Blacks with the disparity widening in early to mid-old age[2,7,10–14].

Reproductive aging has been posited as a vital marker for women's health, particularly for cardio-metabolic health in the midlife period when women are at highest cardio-metabolic risk [15–18]. Menopause is signaled by a women's final menstrual period (FMP) [15–18]. The occurrence and timing of the FMP has been hypothesized to be a potent cardio-metabolic risk factor, as an earlier age at FMP has been associated with increased risk of cardiovascular disease, mortality from ischemic heart disease and stroke[18–21] as well as with risk factors such as low HDL, high waist circumference, and hypertension[22]. Earlier menopause is hypothesized to increase risk by prolonging the time in which a woman experiences low reproductive steroid hormone production[23] potentially causing earlier onset of cardio-metabolic disorders. Menopause also decreases estrogen and progesterone production which contributes to blood vessel stiffening increasing risk of hypertension. These hormonal declines also activate the renin-angiotensin aldosterone system, leading to endothelial dysfunction which increases risk for diabetes and

hypertension[23]. However, evidence is inconsistent on whether there are racial/ethnic differences in the timing of FMP [24–32], a question that necessarily must precede the examination of whether racial/ethnic differences in the timing of FMP could partially explain earlier onset of cardio-metabolic disease and lower life expectancy for Black women.

One possible reason for the inconsistent findings regarding racial/ethnic differences in age at FMP could be selection bias in cohort studies that may differentially affect inclusion of racial/ethnic groups. Selection bias arises from differential probability of selection into (left truncation), and out of (right censoring) the study sample, that may distort study estimates and/or limit generalizability to the populations of interest[33]. Left truncation and right censoring may occur in any cohort - they cause selection bias if the probabilities of left-truncation or right censoring *differ* by population sub-groups of interest, such as by racial/ethnic group. For example, if the recruitment age of a cohort differentially misses subgroups of the population that experience an outcome, e.g., menopause, earlier in life, the study may be differentially left-truncated and would overestimate the average age of FMP in the resulting sample and underestimate racial/ethnic disparities[34]. Left truncation can also be caused by a specific exclusion criterion that differentially affects subgroups of the population. For example, Black women have a much higher rate of surgical menopause (i.e., hysterectomy and/or bilateral oophorectomy) than other racial/ethnic groups in the US[35–37] and women with surgical menopause are frequently excluded at the onset of studies of reproductive aging potentially causing differential left-truncation. Other mechanisms such as non-response bias, or differential willingness/ability to take part in the study once eligible (participation), can also potentially cause differential selection into a cohort[38,39]. Lastly, once the cohort begins differential rates of right censoring can occur due to drop out or excluding certain subgroups in analyses. For example, it is common to right censor women who have experienced surgical menopause in studies of the timing of natural FMP[24,25,27,30,31,35,40]. Given the racial/ethnic differences in the rates of surgical menopause, this practice could cause informative right censoring[41–43], or bias in estimation of racial/ethnic differences in natural FMP due to excluding these subgroups.

Right censoring is a common selection bias addressed in cohorts of aging[43], however, left truncation bias is often relegated to a sentence in the limitations section[34,44]. This study uses the

Study of Women's Health Across the Nation (SWAN) a multi-ethnic cohort of midlife women, that previously found no significant racial/ethnic differences in the timing of FMP (after adjustment)[31], to evaluate the impact of left-truncation and right censoring selection bias on estimates of age at FMP. Exploiting the unique features of SWAN that allow for precise computation of the probability of being selected into the study, this analysis employs a combination of inverse probability weighting[45] and imputation[46] techniques to simultaneously correct for all forms of selection bias and re-estimate racial/ethnic differences in FMP timing, incorporating both surgical and natural menopause to correct for informative right censoring.

3.2 Methods

Data were from the Study of Women's Health Across the Nation (SWAN), a multi-site, multi-ethnic longitudinal cohort (n = 3302) of midlife women, and the screening study used to recruit the cohort (n = 15695) [47]. Each of the 7 sites of SWAN recruited White women, 4 sites recruited Black women (Detroit, MI; Boston, MA; Chicago, IL; Pittsburgh, PA), 1 site recruited Chinese women (Oakland, CA), 1 site Hispanic women (Newark, New Jersey), and 1 site Japanese women (Los Angeles, CA). The eligibility criteria for the screening study were: 40-55 years old, self-identified as the racial/ethnic group for the designated site, spoke either English or the language for the designated site (Spanish in Newark, Japanese in Los Angeles, Cantonese in Oakland), had a primary residence in the designated geographic area for each site and were able to give informed consent[30,47]. Women in the screening study were interviewed via phone or in person to determine if they met the eligibility criteria for the cohort: 42-52 years old, self-identified as the racial/ethnic group for the designated site, primary residence near the site, able to give informed consent, not on hormone therapy (HT) (birth control, fertility drugs, estrogens or progestins, hormone patches or creams, hormone injections or post-menopausal hormones), not currently pregnant, not having had a hysterectomy and/or bilateral oophorectomy, and at least pre to peri menopausal (had a period at least in the last 3 months). The screening study was conducted from 1995-1997 and the cohort, which included in-person clinic visits and interview questions approximately yearly, was conducted from 1995-1997 through 2015-2016. The Institutional Review Boards at each study site approved the protocol and participants provided informed consent for the screening survey and all cohort visits.

Final Menstrual Period (FMP). The primary outcome was age at final menstrual period (FMP), including both surgical and natural FMP within the cohort. Both were determined from annual interviews asking if a woman had experienced amenorrhea for at least 12 months (natural FMP) or had reproductive surgery (bilateral oophorectomy, hysterectomy and/or unilateral oophorectomy). For women that reported both a surgery and natural FMP date, the earlier FMP date was considered their age at FMP. Women were also asked if they were on HT at each visit. If a woman reported HT use and had no subsequent untreated menstrual periods, a natural FMP status could not be determined, and menopausal status was considered missing (right censored).

Racial/Ethnic Group. Self-reported primary racial/ethnic group was collected at baseline. The "Black" category included Black, African-American or those identifying as someone of African origin or descent, the Chinese category included those of Chinese origin or Chinese-Americans, the "Hispanic" category included those identifying as Puerto Rican, Mexican/Mexican-American, Dominican, Central American, Cuban/Cuban-American, South American/Spanish/Other Hispanic, the Japanese category included those of Japanese origin or Japanese-Americans, and the "White" category included those identifying as Caucasian.

Covariates/Mediators. Educational level was assessed at baseline by asking participants their highest level of education achieved (18 categories from did not go to school – doctoral degree), these were collapsed into 3 categories (≤ high school education, some college or an associate degree, ≥ a college degree). Participants were asked, "How hard is it for you to pay for the *very basics* like food, housing, medical care, and heating?", responses were "very hard", "somewhat hard" and "not very hard at all". Participants were asked to rate their general health on a 5-point scale from poor to excellent, answers were collapsed into a 4-level variable from excellent to fair/poor. Participants reported their smoking status including, never a smoker, former smoker, and current smoker. Waist circumference was measured with a measuring tape placed horizontally at the level of the natural waist or the narrowest part of the torso. Body mass index (BMI) was calculated as weight in kilograms divided by measured height in meters squared, measurements were split into categories including less than 25, 25-29.9 and greater than or equal to 30. Self-reported alcohol intake (drinks per week) was collapsed into categories "none to low" (< 2 drinks/week), "moderate" (2-7 drinks/week) and "high" (>7 drinks/week). Physical activity

was assessed using a modified Baecke questionnaire and included sport and exercise activity, non-sport leisure activity, and household and childcare activity[48]: scores ranged from 3 to 15.

Multiple Imputation for Right Censoring. Multiple imputation by chained equations (MICE) was used to impute missing information on the type of FMP (natural or surgical), the age at FMP and covariates/mediators with 50% missing or less in 10 imputation sets[46,49–51]. Predictors for imputed variables included baseline race/ethnicity, study site, age, level of education, employment status, HT use, smoking status, mothers FMP type and age at FMP and parity. Additional predictors included prior-to-FMP diagnoses of fibroids, diabetes, heart attack/angina and cancer (excluding skin cancer). Finally, vasomotor symptoms, menstrual bleeding status, estradiol level, follicle stimulating hormone level, smoking status, alcohol use, BMI, financial hardship, self-reported health, employment status, depression (CES-D), perceived stress level, interpersonal discrimination level[47,52] were obtained from last wave of data collection at which the subject was still known to be menstruating and used a predictors. Participants were missing FMP due to interval censoring (4.91%), HT (16.11%), missing surgery date (0.24%) and loss to follow up (16.38%). Imputation was done in two steps: 1) FMP type (natural or surgical) was imputed and then 2) within each imputation set, FMP age using FMP type as an additional predictor and interaction variable with significant demographic, reproductive and overall health characteristics. Truncated regression (bounded) was used to impute FMP age, with left and right boundaries constructed given the age that FMP was likely to have occurred. Left bounds were based on the last observed menstrual flow without HT-use (with 3-month HT-use washout period) and right bounds were either the first report of 12-months of amenorrhea, first report of stop in HT-use (with 3-month washout period), first report of surgery or age 60 (as 99.7% of women had their FMP by age 60 in the observed sample).

Weighting for Selection into the Cohort. Three inverse probability weights (IPW)[45] were created to account for selection into the cohort to: 1) up-weight women in the cohort who were representative of women who did not make it into the cohort due to ineligibility (eligibility weight), 2) up-weight women who were eligible but unable/unwilling to participate in the cohort (participation weight), and 3) given the 42-52 years age range for recruitment into the study, weight the cohort to represent the cohort that would have been retained had everyone been

recruited at 42 years old (study design weight). Using data from the cross-sectional screening, the inverse odds of eligibility and, given eligibility, participation were calculated. Predictors for each included: race/ethnicity, study site, level of education, marital status, HT use, parity, diagnoses of fibroids, self-reported health, diagnosed heart attack/angina, diagnosed osteoporosis, diagnosed cancer (other than skin cancer), smoking status, BMI with 2-way interactions between race/ethnicity and educational level. Predictors in the final models were selected apriori and using forward-backward selection (accuracy for eligibility was 0.80 and participation 0.65). The study design weight was inverse of the time each woman was eligible over the 10-year period between 42 and 52 years. This is proportional to the probability of each woman, at various ages from 42 to 52, of being excluded from the study given her age and eligibility status (by excluding age as a predictor in the study design weight we avoid "double counting" the effect of age). Weights were multiplied to account for each successive selection mechanism in the order of eligibility, participation, then study design, thus simultaneously accounting for left truncation and selection bias.

Descriptive Statistics and Cox Proportional Hazard Modeling. Mean, proportions and standard errors were calculated and pooled across the imputation sets for all covariates in the total sample and across race/ethnic groups. Bivariate associations were calculated between race/ethnic groups and covariates. The mean, standard errors, and upper/lower confidence intervals were calculated and pooled across imputation sets for each of the selection weights in the cross-sectional screening survey as well as the cohort alone. To determine the potential validity of the weights retained in the cohort, bivariate analyses were performed to determine the difference in weights between the cross-sectional screening and cohort.

Multiple Cox proportional hazard models with age as the timescale were estimated to examine the effect of correcting for the multiple forms of selection bias in the association between racial/ethnic group and age at FMP. Lastly, Cox proportional hazard models and associated median ages were estimated for the model unadjusted for selection and the model fully adjusted for selection further adjusting for potential mediators (Appendix Figure 1). Baseline and time-varying demographic (education and employment), health related (self-reported health, waist circumference, BMI), and health behaviors (smoking status, alcohol use, and physical activity)

were all tested via bivariate analyses and forward-backward selection to determine which would be in the final models. All analyses were conducted in STATA version 16 and R version 4.0.3.

3.3 Results

Table 1 displays descriptive statistics for the cohort at baseline, overall and by race/ethnic group. White women were nearly half of the SWAN sample (47.0%) followed by Black (28.3%), Hispanic (8.7%), Japanese (8.5%) and Chinese (7.6%) women. The average age of the sample was 46.3 years and there was no difference in age between racial/ethnic groups (p-value = 0.176). Generally, 10.1% of the sample had a surgical FMP with Black women having a higher proportion of surgeries (13.8%) and Chinese and Japanese women each having a lower proportion of surgeries (4.8% and 5.7%, respectively, p-value = 0.001). Most of the sample has a college degree or greater (42.6%), however Black women had a higher proportion of women with some college (41.4%) and Hispanic women had a higher proportion of less than or equal to a high school level education (72.1%). Most women had never smoked (58.0%) and only 17.3% were current smokers, however there were higher proportions of current smokers among Black and Hispanic women (24.4% and 18.8%, respectively). The mean waist circumference for the group was 86.4 cm and higher for Black and Hispanic women 93.1 cm and 88.2 cm, respectively. Half of women consumed less than 2 drinks per week (49.9%) with 21.6% consuming >7 drinks per week; however, a larger proportion of White women reported consuming > 7 drinks per week (29.8%).

Table 2 provides results of the Cox proportional hazard models for the association between racial/ethnic group and FMP age, each correcting for an additional form of potential selection bias. In model 1, all racial groups have a higher hazard and earlier natural FMP than White women (range from $HR_{Japanese}$ = 1.02 [0.88-1.19] to $HR_{Hispanic}$ = 1.27 [1.03, 1.57]), only Hispanic women have a significant higher hazard ratio. In model 2, incorporating FMP type as an outcome, all the hazard ratios for natural FMP are lowered and surgical FMP shows a significant earlier FMP for White, Black, Hispanic, and Japanese women compared to White natural FMPs. In model 3 further corrected for right censoring, the hazard ratios for natural FMP increase for Black and Japanese women and are slightly lower for Chinese and Hispanic women while surgical FMP hazard ratios increase for White, Chinese, and Hispanic women and decrease for

Black and Japanese women. In Model 4, further corrected with the eligibility weight, all racial/ethnic groups show a higher hazard of earlier natural FMP that Whites. There is also a higher hazard of earlier surgical FMP for all groups compared to Whites with natural FMP. In Model 5 further corrected with the participation weight, all the hazard ratios for each racial/ethnic group lowered for natural FMP except for Black women. For surgical FMP, three hazard ratios lowered from the last model while Chinese and Hispanic women increased. In the last model (model 6) further incorporating the study design weight, all racial/ethnic groups had a higher hazard of earlier natural FMP compared to Whites except Japanese women (HR_{Black} = 1.15 [1.04, 1.27], $HR_{Chinese}$ = 1.02 [0.87, 1.16], $HR_{Hispanic}$ = 1.18 [0.97, 1.38] and $HR_{Japanese}$ = 0.97 [0.83, 1.10]). All racial/ethnic groups had a higher hazard of earlier surgical FMP compared to Whites with natural FMP (range from $HR_{Chinese}$ = 1.57 [0.71, 2.43] to HR_{Black} = 3.02 [2.58, 3.45]). Figure 2 shows the estimated distribution for natural and surgical FMP after fully adjusting for selection which shows similar trends to Model 6.

Table 3 shows Cox proportional hazard models and associated median ages adjusting for covariates in the model not accounting for selection (model 1) and in the model fully accounting for selection (model 6). In the model not accounting for selection (model 1), all the hazard ratios lowered and become non-significant, except for Japanese women who had a later natural FMP than White women ($HR_{Japanese}$ = 0.84 [0.71, 1.00], $Median_{Japanese}$ = 53.0 [52.5, 53.3] compared to $Median_{White}$ = 52.4 [52.1, 52.6]). In the model fully accounting for selection (model 6), Black women show a significant earlier natural FMP and Japanese women a later natural FMP compared to White women (HR_{Black} = 1.13 [1.00, 1.26] with $Median_{Black}$ = 51.5 [51.3, 51.6] and $HR_{Japanese}$ = 0.83 [0.69, 0.98] with $Median_{Japanese}$ = 52.4 [52.1, 52.8] compared to $Median_{White}$ = 51.9 [51.8, 52.1]). For surgical FMP, all groups other than Chinese women (n_{events} = 13) had a significant earlier surgical FMP than White women with natural FMP (range from $HRJ_{apanese}$ = 1.43 [1.02, 1.85] with $Median_{Japanese}$ = 51.4 [50.7, 51.7] to HR_{Black} = 3.21 [2.80, 3.62] with $Median_{Black}$ = 47.2 [46.9, 47.6]) compared to $Median_{White}$ = 51.9 [51.8, 52.1]).

3.4 Discussion

This is the first study, to our knowledge, to simultaneously correct for left-truncation and right censoring in a cohort of aging - examining how racial/ethnic differences in a health outcome are

affected. Results demonstrate that estimates of reproductive aging that ignore selection biases led to falsely high estimations of the average FMP age and underestimation of race/ethnic differences in menopausal timing. These biases particularly effected estimates of menopausal age in Black women whose age of natural FMP was overestimated in the unaccounted model by an average of 1.1 years. After selection biases stemming from selection into the cohort and informative censoring were accounted for in the SWAN study, Black women had an earlier menopause compared to White women by an average of 1.2 years – 0.60 years earlier for natural menopause and 1.8 years earlier for surgical menopause. These Black –White differences remained after controlling for socioeconomic indicators, overall health, and health behaviors signaling that structural factors contributing to the "weathering"[7,8] process may be at play.

Selection bias into and out of the SWAN cohort lead to an underestimate of racial/ethnic differences in FMP, after accounting for selection racial/ethnic differences were present. In a simulation study by Mayeda et al., researchers found that survival bias, a type of left-truncation caused by mortality prior to study entry, can cause underestimation of the association between education and cognitive decline[44]. Like Mayeda et al., this study showed that racial/ethnic differences in menopausal timing are also underestimated when left truncation is unaccounted for. There was an average positive bias of 0.66 years between the "biased" and "unbiased" models in this study, however the magnitude varied by racial/ethnic group where Black women had the highest amount of bias and Chinese women had the lowest amount of bias (+1.10 and +0.40 years, respectively). In a simulation analysis by Cain et al. that estimated natural FMP, when a cohort was started at 42.5-52.5 years old the effect of left truncation on estimations of natural menopausal timing resulted in a positive bias of 1.29 years[34]. SWAN started at 42-52 years old and the bias in this study are lower than what was predicted in the Cain et al. article. This could be partially due to the difference in outcome (natural FMP versus overall FMP in this study) however, the amount of bias is similar for Black women. Further research is needed to disentangle the cause of the difference, however, the differential bias for Black women point to differential left truncation between racial/ethnic groups – an important added element to consider when assessing left truncation.

Unlike the Cain et al. and Mayeda et al. studies[34,44] that quantified bias in simulations where the departure of the biased estimate from the actual estimate can be observed, this study is a "real world" example of correcting for selection. Although results are similar to the simulations[34,44], the eligibility and participation weights are operating on the assumption that at least some women retained in the SWAN cohort have similar odds of being eligible or participating to those who were left-truncated. However, the weights in the cohort are significantly smaller than those in the cross-sectional study (appendix table 1 and figure 2), reflecting that some of the women with lower odds of inclusion (and thus higher weights) are not retained in the cohort to be upweighted. This is true for both weights, especially the eligibility weight, and occurs equally across racial/ethnic groups. Highlighting that although weights can mitigate some of the bias due to left truncation, they may not mitigate all the bias that would be addressed by modifying the study design to include those who had reached FMP upon recruitment[34]. This also underlines the usefulness of the study design weight, an IPW weight that uses information from the cohort rather than depending on representation in the cohort of those with lower odds of inclusion. Although it does not address left-truncation completely, results show that weighting women in the cohort as if they were recruited at age 42 years with the study design weight contributes the most to addressing bias compared to the eligibility and participation weights alone.

Results show that Black women had an earlier natural menopause than White women robust to adjustment for potential mediating factors. In contrast to other studies[25–28,32], previous longitudinal analysis in SWAN had not reported Black/White differences in the timing of FMP after adjustment[25]. After accounting for selection biases Black women had an earlier natural FMP compared to White women by 0.4 years and, although only marginally significant, Hispanic women followed a similar trend of an earlier natural FMP by 0.5 years. Other studies conducted in younger age groups, most of which censored women with surgical menopause have found as large as 1 year difference in median age of natural FMP[24,26,30,53] in Black/Hispanic women compared to Whites[24–27,30,31,40,53]. It is common in reproductive aging studies to censor women who have had surgical menopause[24,25,27,30,31,35,40], instead these analyses incorporated surgical menopause as a moderator and used a novel two step approach to multiple imputation. Those loss to follow up were not assumed to have natural menopause and type of FMP was used as an interaction variable in imputation prediction models to account for potential differences in

FMP timing by FMP type. This procedure allowed an examination of potential informative censoring by surgical menopause. When surgical menopause was censored the hazard for Black and Hispanic women went up indicating that surgical menopause may be a form of informative censoring or censoring that is related to the FMP process[41,42] and differential by racial/ethnic group causing bias in estimation. Additionally, attenuation of racial/ethnic differences in FMP by socioeconomic status, health status and health behaviors were substantial in the "biased" model but minimal in the "unbiased" models. The retained SWAN cohort tended to be of higher socioeconomic status and healthier than women originally screened for the study (Aim 1.1), thus in upweighting the women representative of those screened out of the cohort there was potentially more variability introduced in these variables leading to better estimates of the true effect of these factors. Results on the non-attenuation of racial/ethnic differences in FMP in the "unbiased" models were similar to other longitudinal studies with a younger age of entry[29]. It could be that there are other aspects of the "weathering"[54] process not captured by these indicators, that are structural[9,55] or interpersonal. For example, Bromberger et al. found that disparities in natural FMP were increased for Black women who reported an ongoing psychosocial stress[26].

Although White women had a slightly higher prevalence of surgical menopause overall, Black and Hispanic women had surgical menopause at significantly earlier ages. Black women on average had reproductive surgeries 2 years earlier and Hispanic women 1 year earlier than White women. The largest proportion of left truncation in the SWAN study stemmed from Black women excluded due to surgical menopause (30% [Aim 1.1]). Previous studies have highlighted the high prevalence of surgical menopause among Black and Hispanic women, especially Black women[35–37]. The differential risk of hysterectomy/oophorectomy could be partially due to the higher prevalence of reproductive morbidities (such as uterine fibroids) earlier in life for Black women, which are often treated with surgery[56–58]. However similar to the results here that were robust to adjustment, Powell et al. also noted the lack of change in risk of surgical menopause after controlling for known risk factors such as socioeconomic status and obesity[36]. Suggesting that the higher prevalence, independent of known risk factors, is consistent with an "overuse of elective hysterectomy" in these populations[36].

This study has some limitations. The SWAN study recruited the Hispanic, Chinese and Japanese racial/ethnic groups at one site each, thus geographic/site differences cannot be disentangled from racial/ethnic differences for these groups. Furthermore, the low prevalence of surgical FMPs in Chinese, Hispanic and Japanese women produced imprecise estimates of age at surgical menopause for these populations. Lastly, the hazard ratios in our "biased" model differs slightly from the hazard ratio estimates in previous SWAN longitudinal analyses of natural FMP[31]. However, the previous analyses were conducted in 2007 and included fewer observed FMPs (1403 natural FMP events in original analyses[31] versus 1804 in this analyses).

In conclusion, although mitigation may be imperfect for the reasons stated above, we have shown that not considering left truncation biases estimates of racial disparities in FMP. Furthermore, censoring for surgical FMP biases the estimates of natural FMP – a common method in reproductive aging research. Results show that Black women have an earlier natural and surgical FMP than White women that is robust to adjustment for known risk factors, a finding with potential implications for Black women's overall aging and cardio-metabolic health. Results may be driven by a combination of "weathering" and overuse of elective surgeries like hysterectomy. Further research is needed to clarify the causes and consequences of earlier reproductive aging among US Black women.

Table 3.1 Baseline Characteristics of the Multiply Imputed Data in SWAN (n = 3302)

	Total	**White** (n = 1551)	**Black** (n = 935)	**Chinese** (n = 250)	**Hispanic** (n = 286)	**Japanese** (n = 281)	p-value[b]
Racial/Ethnic Group[a]	-	47.0	28.3	7.6	8.7	8.5	-
Age (Mean, SE)	46.3 (0.05)	46.3 (0.07)	46.2 (0.09)	46.5 (0.16)	46.3 (0.16)	46.7 (0.16)	0.176
Final Menstrual Period Type							0.001
Natural	89.9 (0.01)	90.5 (0.01)	86.2 (0.01)	95.2 (0.01)	89.4 (0.03)	94.3 (0.01)	
Surgical	10.1 (0.01)	9.5 (0.01)	13.8 (0.01)	4.8 (0.01)	10.6 (0.03)	5.7 (0.01)	
Educational Level							0.000
<= high school	25.1 (0.01)	16.1 (0.01)	26.8 (0.01)	29.0 (0.03)	72.1 (0.03)	18.3 (0.02)	
some college	32.3 (0.01)	30.6 (0.02)	41.4 (0.02)	21.8 (0.02)	18.8 (0.02)	34.3 (0.02)	
>= college	42.6 (0.01)	53.3 (0.02)	31.9 (0.02)	49.2 (0.02)	9.1 (0.02)	47.4 (0.02)	
Financial Hardship							0.000
very hard	9.3 (0.01)	6.0 (0.01)	12.5 (0.01)	5.2 (0.01)	26.4 (0.03)	3.6 (0.01)	
somewhat hard	30.7 (0.01)	26.2 (0.01)	33.8 (0.02)	22.9 (0.03)	55.1 (0.03)	26.7 (0.03)	
not very hard	60.0 (0.01)	67.8 (0.01)	53.6 (0.02)	71.9 (0.03)	18.5 (0.02)	69.8 (0.03)	
Self-Reported Health							0.000
excellent	21.3 (0.01)	29.2 (0.01)	15.1 (0.01)	16.8 (0.02)	4.9 (0.01)	19.3 (0.02)	
very good	36.3 (0.01)	42.2 (0.01)	32.8 (0.02)	29.4 (0.03)	21.7 (0.02)	36.8 (0.03)	
good	29.2 (0.01)	22.1 (0.01)	35.7 (0.02)	32.4 (0.03)	46.2 (0.03)	26.1 (0.03)	
fair/poor	13.2 (0.01)	6.5 (0.01)	16.4 (0.01)	21.4 (0.03)	27.1 (0.03)	17.8 (0.02)	
Smoking Status							0.000
never	58.0 (0.01)	51.7 (0.01)	52.9 (0.02)	94.4 (0.02)	66.8 (0.03)	68.6 (0.03)	
former	24.7 (0.01)	32.2 (0.01)	22.7 (0.01)	3.4 (0.01)	14.4 (0.02)	20.0 (0.02)	
current	17.3 (0.01)	16.2 (0.01)	24.4 (0.01)	2.2 (0.01)	18.8 (0.02)	11.4 (0.02)	
Waist Circumference (cm)	86.4 (0.28)	85.7 (0.41)	93.1 (0.54)	77.3 (0.65)	88.2 (0.83)	73.5 (0.52)	0.000
Body Mass Index							0.000
< 25	39.8 (0.01)	42.8 (0.01)	18.3 (0.01)	76.4 (0.03)	22.8 (0.02)	79.0 (0.02)	
25-29.9	26.2 (0.01)	25.3 (0.01)	28.9 (0.01)	18.4 (0.02)	39.6 (0.03)	15.7 (0.02)	
>= 30	34.0 (0.01)	31.9 (0.01)	52.8 (0.02)	5.2 (0.01)	37.6 (0.03)	5.3 (0.01)	
Alcohol Consumption (servings/week)							0.000
none/low (< 2)	49.9 (0.01)	39.6 (0.01)	57.0 (0.02)	79.1 (0.03)	50.7 (0.03)	56.1 (0.03)	
moderate (2-7)	28.6 (0.01)	30.6 (0.01)	26.6 (0.01)	14.9 (0.02)	41.6 (0.03)	22.8 (0.03)	
high (>7)	21.6 (0.01)	29.8 (0.01)	16.4 (0.01)	6.0 (0.02)	7.7 (0.02)	21.1 (0.02)	
Physical Activity Score*	7.6 (0.03)	8.0 (0.05)	7.3 (0.06)	7.3 (0.11)	6.8 (0.09)	7.9 (0.10)	0.000

Note: [a]Racial/Ethnic group was not imputed. [b]Difference between racial/ethnic groups.

Table 3.2 Cox Proportional Hazard Models for Age at Final Menstrual Period (FMP) Accounting for Successive Selection Mechanisms

	Model 1[a] *Unadjusted for Selection* n_{events} = 1804				Model 2[b] *Incorporating FMP Type* n_{events} = 2059				Model 3[c] *+ Imputed for Right Censoring* n_{events} = 3302				Model 4[d] *+ IPW for Eligibility* n_{events} = 3302				Model 5[e] *+ IPW for Participation* n_{events} = 3302				Model 6[f] *Fully Adjusted for Selection* n_{events} = 3302			
	Hazard	LCI	UCI	p-value	Hazard	LCI	UCI	p-value	Hazard	LCI	UCI	p-value	Hazard	LCI	UCI	p-value	Hazard	LCI	UCI	p-value	Hazard	LCI	UCI	p-value
Racial Group x FMP Type (ref = White, Natural FMP)																								
Natural FMP																								
Black	1.09	0.97	1.22	0.131	1.05	0.94	1.18	0.367	1.11	1.01	1.21	0.023	1.12	1.01	1.23	0.025	1.12	1.01	1.24	0.023	1.15	1.04	1.27	0.008
Chinese	1.08	0.92	1.28	0.336	1.04	0.88	1.23	0.634	1.03	0.89	1.17	0.342	1.06	0.91	1.21	0.237	1.03	0.87	1.18	0.368	1.02	0.87	1.16	0.411
Hispanic	1.27	1.03	1.57	0.026	1.25	1.01	1.54	0.037	1.19	1.04	1.34	0.012	1.10	0.92	1.29	0.147	1.06	0.88	1.24	0.266	1.18	0.97	1.38	0.062
Japanese	1.02	0.88	1.19	0.793	0.95	0.81	1.11	0.498	0.97	0.83	1.10	0.690	1.01	0.88	1.13	0.459	0.98	0.86	1.11	0.608	0.97	0.83	1.10	0.683
Surgical FMP																								
White					1.25	1.03	1.52	0.025	1.34	1.15	1.53	0.002	1.33	1.07	1.59	0.016	1.24	0.90	1.57	0.105	1.58	1.24	1.93	0.005
Black					2.59	2.10	3.19	0.000	2.41	2.20	2.63	0.000	2.25	1.85	2.64	0.000	2.05	1.58	2.53	0.002	3.02	2.58	3.45	0.000
Chinese[g]					0.94	0.52	1.70	0.827	0.97	0.39	1.55	0.542	1.07	0.28	1.86	0.432	1.12	0.34	1.90	0.390	1.57	0.71	2.43	0.151
Hispanic[g]					1.83	1.10	3.06	0.021	2.07	1.60	2.55	0.001	2.15	1.41	2.88	0.021	2.37	1.68	3.06	0.007	2.63	2.02	3.23	0.001
Japanese[g]					1.88	1.11	3.20	0.019	1.27	0.77	1.77	0.176	1.51	0.98	2.04	0.063	1.49	0.98	2.00	0.063	1.43	0.99	1.87	0.055

Note: Each subsequent model has the "corrected" features of the last with additional corrections, denoted by + sign. Inverse Probability Weight (IPW).

[a]Model 1 - right censored for any missing FMPs and for observed surgical FMPs.

[b]Model 2 - only censors for unobserved FMPs and includes in observed surgical FMPs by incorporating an interaction term between racial/ethnic group and FMP type.

[c]Model 3 - additionally imputes data pooled across 10 imputed sets removing any right censoring on FMP.

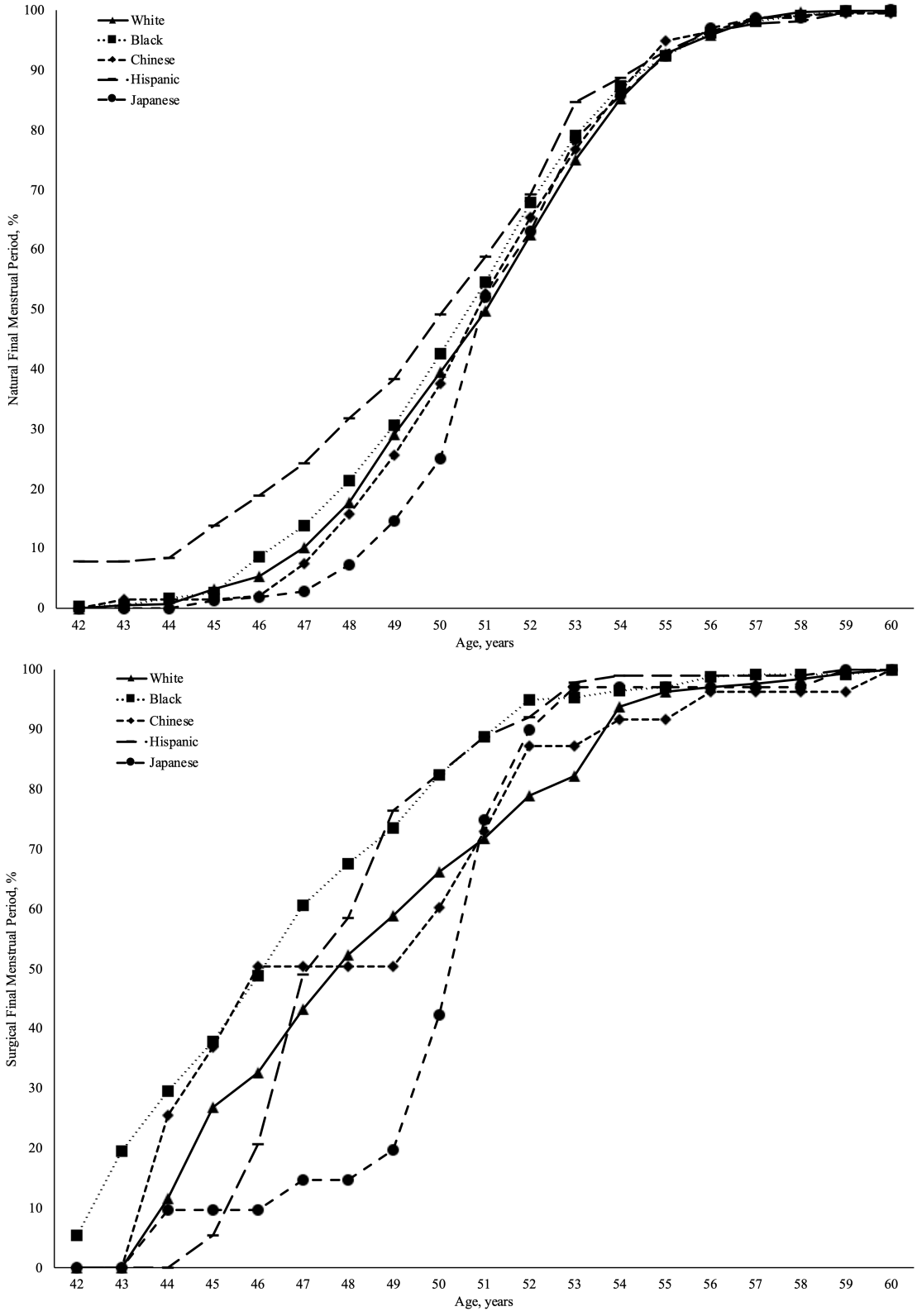
[d]Model 4 - additionally incorporates a weight for eligibility into the study.

[e]Model 5 - additionally incorporates a weight for participation.

[f]Model 6 - additionally incorporates weight for study design.

[g]Chinese (n = 13), Hispanic (n = 39), Japanese (n = 16), all other groups > 90 observations.

Figure 3.1 Estimated Distribution of Age at Final Menstrual Period (Natural and Surgical) by Racial/Ethnic Group after Fully Adjusting for Selection



Note: Distribution in one imputation set ($SE_{between}$ = 0.003).

Table 3.3 Adjusted Cox Proportional Hazard Models and Medians for Age at Final Menstrual Period

	Adjusted Model 1[ab] *Unadjusted for Selection*					**Adjusted Model 6[cb]** *Fully Adjusted for Selection*				
	Hazard	LCI	UCI	p-value	Median Age	Hazard	LCI	UCI	p-value	Median Age
Racial Group x FMP Type (ref = White, Natural FMP)										
Natural FMP										
White	*reference*				52.4 (52.1, 52.6)	*reference*				51.9 (51.8, 52.1)
Black	0.98	0.86	1.11	0.73	52.6 (52.4, 52.9)	1.13	1.00	1.26	0.04	51.5 (51.3, 51.6)
Chinese	0.88	0.73	1.07	0.21	52.7 (52.2, 53.5)	0.94	0.76	1.13	0.73	52.0 (51.6, 52.4)
Hispanic	1.04	0.81	1.33	0.77	52.1 (51.4, 53.1)	1.03	0.79	1.28	0.39	51.4 (51.2, 51.8)
Japanese	0.84	0.71	1.00	0.05	53.0 (52.5, 53.3)	0.83	0.69	0.98	0.99	52.4 (52.1, 52.8)
Surgical FMP										
White						1.58	1.19	1.98	0.01	48.9 (48.7, 49.4)
Black						3.21	2.80	3.62	0.00	47.2 (46.9, 47.6)
Chinese[d]						1.29	0.42	2.15	0.28	50.5 (46.8, 52.2)
Hispanic[d]						2.30	1.64	2.96	0.01	48.3 (47.9, 49.7)
Japanese[d]						1.43	1.02	1.85	0.04	51.4 (50.7, 51.7)

Note: [a]Model 1 - right censored for any missing FMPs and for observed surgical FMPs (n_{events} = 1653).
[b]Adjusted for: baseline education, baseline self-reported health, baseline and last known prior to FMP waist circumference, baseline smoking status, baseline alcohol use, baseline physical activity score.
[c]Model 6 - adjusts for all forms of selection by incorporating imputation and weights (n_{events} = 3302).
[d]Chinese (n = 13), Hispanic (n = 39), Japanese (n = 16), all other groups > 90 observations.

Figure 3.2 Analytic Model for the Association between Racial/Ethnic Group and Age at Final Menstrual Period (FMP) [appendix]

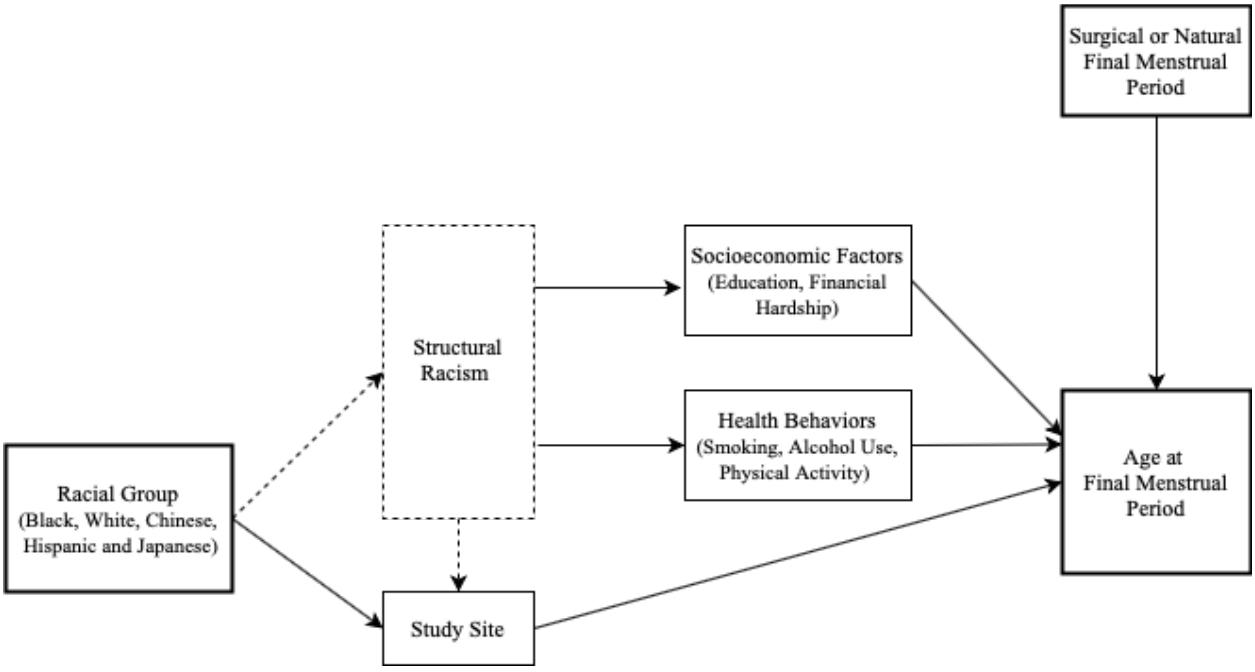


Table 3.4 Descriptive Statistics for Weights by Racial/Ethnic Group in Cross-Sectional Screening and Cohort [appendix]

		In Cross-Sectional + Cohort				In Cohort Only				
		Mean (SE)	LCI	UCI	p-value[a]	Mean (SE)	LCI	UCI	p-value[a]	p-value[b]
Eligibility Weight	Total	2.84 (0.034)	2.77	2.90	-	1.65 (0.016)	1.62	1.68	-	0.000
	Racial/Ethnic Group									
	White	2.89 (0.052)	2.79	2.99	-	1.66 (0.026)	1.61	1.71	-	0.000
	Black	3.10 (0.057)	2.99	3.21	0.009	1.82 (0.030)	1.76	1.88	0.000	0.000
	Chinese	1.49 (0.036)	1.42	1.56	0.000	1.30 (0.018)	1.26	1.33	0.000	0.000
	Hispanic	2.96 (0.106)	2.76	3.17	0.493	1.66 (0.047)	1.57	1.76	0.995	0.000
	Japanese	1.64 (0.046)	1.55	1.73	0.000	1.30 (0.028)	1.24	1.35	0.000	0.000
Participation Weight	Total	2.45 (0.012)	2.43	2.47	-	1.94 (0.015)	1.91	1.97	-	0.000
	Racial/Ethnic Group									
	White	2.69 (0.021)	2.65	2.73	-	2.03 (0.026)	1.98	2.08	-	0.000
	Black	2.03 (0.011)	2.01	2.05	0.000	1.82 (0.019)	1.78	1.86	0.000	0.000
	Chinese	1.40 (0.008)	1.38	1.41	0.000	1.38 (0.011)	1.36	1.41	0.000	0.195
	Hispanic	3.14 (0.035)	3.07	3.21	0.000	2.75 (0.055)	2.64	2.86	0.000	0.000
	Japanese	1.58 (0.006)	1.56	1.59	0.000	1.55 (0.010)	1.53	1.57	0.000	0.004
Study Design Weight	Total	-				1.41 (0.031)	1.35	1.47	-	
	Racial/Ethnic Group									
	White					1.35 (0.031)	1.29	1.41	-	
	Black					1.52 (0.073)	1.38	1.67	0.023	
	Chinese	-				1.24 (0.035)	1.17	1.31	0.353	
	Hispanic					1.58 (0.152)	1.28	1.87	0.048	
	Japanese					1.30 (0.140)	1.02	1.57	0.616	

Note: [a]Significance for difference from White (reference).
[b]Significance for difference between cross-sectional and cohort weights.

3.5 References

1. Hayward, M. D., Crimmins, E. M., Miles, T. P. & Yang, Y. U. The Significance of Socioeconomic Status in Explaining the Racial Gap in Chronic Health Conditions Author (s): Mark D . Hayward , Toni P . Miles , Eileen M . Crimmins and Yu Yang Published by : American Sociological Association Stable URL : http://www.j. **65**, 910–930 (2016).
2. Levine, M. E. & Crimmins, E. M. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc. Sci. Med.* **118**, 27–32 (2014).
3. Kochanek, K. D., Arias, E. & Anderson, R. N. *How Did Cause of Death Contribute to Racial Differences in Life Expectancy in the United States in 2010? NCHS data brief, no 125* (2013).
4. Crimmins, E. M. & Saito, Y. Trends in healthy life expectancy in the United States, 1970–1990: gender, racial, and educational differences. *Soc. Sci. Med.* **52**, 1629–1641 (2001).
5. Olshansky, S. J. *et al.* Differences In Life Expectancy Due To Race And Educational Differences Are Widening, And Many May Not Catch Up. *Health Aff.* **31**, 1803–1813 (2012).
6. Wong, M. D., Shapiro, M. F., Boscardin, W. J. & Ettner, S. L. Contribution of Major Diseases to Disparities in Mortality. *N. Engl. J. Med.* **347**, 1585–1592 (2002).
7. Geronimus, A. T., Hicken, M., Keene, D. & Bound, J. 'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am. J. Public Health* **96**, 826–833 (2006).
8. Geronimus, A. T., Bound, J., Keene, D. & Hicken, M. Black-White Differences in Age Trajectories of Hypertension Prevalence Among Adult Women and Men, 1999-2002. *Ethn. Dis.* **17**, 40–48 (2007).
9. Phelan, J. C. & Link, B. G. Is Racism a Fundamental Cause of Inequalities in Health? doi:10.1146/annurev-soc-073014-112305
10. Nan, |, Astone, M., Ensminger, M. & Juon, H. S. *Early Adult Characteristics and Mortality Among Inner-City African-American Women*. *American Journal of Public Health* **92**, (2002).
11. Geronimus, A. T., Bound, J., Waidmann, T. A., Hillemeier, M. M. & Burns, P. B. Excess Mortality among Blacks and Whites in the United States. *N. Engl. J. Med.* **335**, 1552–1558 (1996).
12. Geronimus, A. T., Bound, J., Waidmann, T. A., Colen, C. G. & Steffick, D. Inequality in Life Expectancy, Funcational Status, and Active Life Expectancy Across Selected Black and White Populations in the United States. *Demography* **38**, 227–251 (2001).
13. Elo, I. T. & Preston, S. H. *Educational Differential in Mortality: United States*. *Soc. Sci. Med* **42**, (1996).
14. Geronimus, A. T. *et al.* Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-based Sample. *J. Health Soc. Behav.* **56**, 199–224 (2015).
15. Snowdon, D. A. *et al.* Is Early Natural Menopause a Biologic Marker of Health and Aging? *Am. J. Public Health* **79**, 709–714 (1989).
16. Wise, P. M., Krajnak, K. M. & Kashon, M. L. Menopause: The Aging of Multiple Pacemakers. *Am. Assoc. Adv. Sci.* **273**, 67–70 (1996).
17. Cooper, G. S. & Sandler, D. P. Age at Natural Menopause and Mortality. *Ann. Epidemiol.*

8, 229–235 (1998).
18. Gold, E. B. The timing of the age at which natural menopause occurs. *Obstet. Gynecol. Clin. North Am.* **38**, 425–40 (2011).
19. Appelman, Y., van Rijn, B. B., ten Haaf, M. E., Boersma, E. & Peters, S. A. E. Sex differences in cardiovascular risk factors and disease prevention. *Atherosclerosis* **241**, 211–218 (2015).
20. Muka, T. *et al.* Association of Age at Onset of Menopause and Time Since Onset of Menopause With Cardiovascular Outcomes, Intermediate Vascular Traits, and All-Cause Mortality. *JAMA Cardiol.* **1**, 767 (2016).
21. Yoshida, Y. *et al.* Early Menopause and Cardiovascular Disease Risk in Women With Or Without Type 2 Diabetes: A Pooled Analysis of 9,374 Postmenopausal Women. *Diabetes Care* **70**, 992-P (2021).
22. Bleil, M. E., Gregorich, S. E., Mcconnell, D., Rosen, M. P. & Cedars, M. I. DOES ACCELERATED REPRODUCTIVE AGING UNDERLIE PRE-MENOPAUSAL RISK FOR CARDIOVASCULAR DISEASE? *Menopause* **20**, 1139–1146 (2013).
23. Mendelsohn, M. E. & Karas, R. H. The Protective Effects of Estrogen on the Cardiovascular System. *N. Engl. J. Med.* **340**, 1801–1811 (1999).
24. MacMahon, B. & Worcester, J. *Age at menopause. United States--1960-1962. Vital and health statistics. Series 11, Data from the National Health Survey* (1966).
25. Luborsky, J. L., Meyer, P., Sowers, M. F., Gold, E. B. & Santoro, N. Premature menopause in a multi-ethnic population study of the menopause transition*. *Hum. Reprod.* **18**, 199–206 (2002).
26. Bromberger, J. T. *et al.* Prospective Study of the Determinants of Age at Menopause. *Am. J. Epidemiol. Copyr. O* **145**, (1997).
27. Stanford, J. L., Hartge, P., Brinton, L. A., Hoover', R. N. & Brookmeyer, R. Factors Influencing the Age of Natural Menopause. *J. Chronic Dis.* **40**, 995–1002 (1987).
28. Butts, S. F. & Seifer, D. B. Racial and ethnic differences in reproductive potential across the life cycle. *Fertil. Steril.* **93**, 681–690 (2010).
29. Delellis Henderson, K., Bernstein, L., Henderson, B., Kolonel, L. & Pike, M. C. Original Contribution Predictors of the Timing of Natural Menopause in the Multiethnic Cohort Study. *Am. J. Epidemiol.* **167**, 1287–1294 (2008).
30. Gold, E. B. *et al.* Factors Associated with Age at Natural Menopause in a Multiethnic Sample of Midlife Women. *Am. J. Epidemiol.* **153**, 865–874 (2001).
31. Gold, E. B. *et al.* Factors related to age at natural menopause: Longitudinal analyses from SWAN. *Am. J. Epidemiol.* (2013). doi:10.1093/aje/kws421
32. Appiah, D., Nwabuo, C. C., Ebong, I. A., Wellons, M. F. & Winters, S. J. Trends in Age at Natural Menopause and Reproductive Life SpanAmongUSWomen, 1959-2018. *JAMA - J. Am. Med. Assoc.* **325**, 1328–1330 (2021).
33. Rothman, K. J., Greenland, S. & Lash, T. J. *Modern Epidemiology*. (Lippincott Williams and Wilkins, 2014).
34. Cain, K. C. *et al.* Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes. *Am. J. Epidemiol.* **173**, 1078–1084 (2011).
35. WRIGHT, M. A. *et al.* Changing trends in Black-White racial differences in surgical menopause: a population-based study. *Am. J. Obstet. Gynecol.* (2021). doi:10.1016/j.ajog.2021.05.045
36. Powell, L. H. *et al.* Ethnic differences in past hysterectomy for benign conditions.

Women's Heal. Issues **15**, 179–186 (2005).
37. Bower, J. K., Schreiner, P. J., Sternfeld, B. & Lewis, C. E. Black-white differences in hysterectomy prevalence: The CARDIA study. *Am. J. Public Health* **99**, 300–307 (2009).
38. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
39. Nohr, E. A., Frydenberg, M., Henriksen, T. B. & Olsen, J. Does low participation in cohort studies induce bias? *Epidemiology* **17**, 413–418 (2006).
40. Merrill, R. M., Layman, A. B., Oderda, G. & Asche, C. Risk Estimates of Hysterectomy and Selected Conditions Commonly Treated with Hysterectomy. *Ann. Epidemiol.* **18**, 253–260 (2008).
41. Moore, D. F. *Applied Survival Analysis Using R*. (Springer, 2016). doi:10.1007/978-3-319-31245-3
42. Siannis, F., Copas, J. & Lu, G. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics* **6**, 77–91 (2005).
43. Howe, C. J., Cole, S. R., Lau, B., Napravnik, S. & Eron, J. J. Selection bias due to loss to follow up in cohort studies. *Epidemiology* **176**, 139–148 (2017).
44. Mayeda, E. R., Filshtein, T. J., Tripodis, Y., Maria Glymour, M. & Gross, A. L. Does selective survival before study enrolment attenuate estimated effects of education on rate of cognitive decline in older adults? A simulation approach for quantifying survival bias in life course epidemiology. *Int. J. Epidemiol.* **47**, 1507–1517 (2018).
45. Seaman, S. R. & White, I. R. Review of inverse probability weighting for dealing with missing data. doi:10.1177/0962280210395740
46. Rubin, D. B. Multiple Imputations in Sample Surveys - A Phenomenological Baysian Approach to Nonresponse. *Proc. Surv. Res. methods Sect. Am. Stat. Assoc. Am. Stat. Assoc.* **1**, 20–34 (1978).
47. Sowers, M. *et al.* SWAN: A Multicenter, Multiethnic, Community Based Cohort Study of Women and the Menopausal Transition. in *Menopause: Biology and Pathobiology* 175–188 (2000).
48. Baecke, J. Determinants of Body Fatness in Young Adults Living in a Dutch Community. *Doctoral Dissertation Baecke* (1982).
49. Marshall, A., Altman, D. G., Royston, P. & Holder, R. L. Comparison of techniques for handling missing covariate data within prognostic modelling studies: A simulation study. *BMC Med. Res. Methodol.* **10**, 1–16 (2010).
50. Lee, K. J. & Carlin, J. B. Recovery of information from multiple imputation: A simulation study. *Emerg. Themes Epidemiol.* **9**, 1–10 (2012).
51. Lee, K. J., Roberts, G., Doyle, L. W., Anderson, P. J. & Carlin, J. B. Multiple imputation for missing data in a longitudinal cohort study: a tutorial based on a detailed case study involving imputation of missing outcome data. *Int. J. Soc. Res. Methodol.* **19**, 575–591 (2016).
52. Lewis, T. T., Yang, F. M., Jacobs, E. A. & Fitchett, G. Racial/ethnic differences in responses to the everyday discrimination scale: A differential item functioning analysis. *Am. J. Epidemiol.* **175**, 391–401 (2012).
53. Stanford, J. L., Hartge, P., Brinton, L. A., Hoover, R. N. & Brookmeyer, R. Factors influencing the age at natural menopause. *J. Chronic Dis.* **40**, 995–1002 (1987).
54. Geronimus, A. T. *et al.* Do US black women experience stress-related accelerated

biological aging?: A novel theory and first population-based test of black-white differences in telomere length. *Hum. Nat.* **21**, 19–38 (2010).
55. Williams, D. R. & Collins, C. Racial Residential Segregation: A Fundamental Cause of Racial Disparities in Health. *Public Health Rep.* **116**, (2001).
56. Baird, D. D., Dunson, D. B., Hill, M. C., Cousins, D. & Schectman, J. M. High cumulative incidence of uterine leiomyoma in black and white women: Ultrasound evidence Donna. *Am. J. Obstet. Gynecol.* **188**, 100–107 (2003).
57. Marsh, E. E. *et al.* Racial differences in fibroid prevalence and ultrasound findings in asymptomatic young women (18-30 years old): a pilot study. *Fertil. Steril.* **99**, 1951–1957 (2013).
58. Okolo, S., Medical Executive Director, F. & of Ob-Gyn, P. Incidence, aetiology and epidemiology of uterine fibroids. *Best Pract. Res. Clin. Obstet. Gynaecol.* **22**, 571–588 (2008).

Chapter 4 The Influence of Multiple Sources of Selection Bias on Racial Differences in Cardio-Metabolic Disease Onset in a Cohort of Midlife Aging

4.1 Introduction

Selection bias is a distortion in estimation caused by differential selection into or out of a study or analyses[1,2]. It can arise from multiple sources in a cohort study. Cohort studies typically select individuals at a life stage where they are free from the outcome or disease of interest and follow participants over time to observe outcomes of interest and the factors that influence onset of outcomes[1,3]. Selection processes at study commencement cause two forms of selection bias, left truncation and left censoring. Left truncation occurs when individuals who have already had the outcome of interest are not included in the study. Left censoring occurs when individuals who have had the outcome of interest are included in the study. Left truncated individuals are excluded from any analyses as they are not in the cohort while left censored individuals, although retained in the cohort, are frequently excluded from analyses because they have the outcome at study baseline and thus the factors that influence onset were not observed. Lastly, right censoring occurs when the outcome of interest is not observed for a participant in a cohort study[3], either from non-occurrence of the outcome prior to study termination, from loss to follow up, or due to missing variables needed to determine the outcome of interest. Bias can occur via any of these three mechanisms if the probability of selection differs across relevant characteristics that influence causal estimation, or the association between exposure and outcome of interest[4–7].

Given persistent large racial disparities in health[8–11] it is probable that selection via these three mechanisms could affect racial/ethnic groups differentially biasing estimation. Cardio-metabolic diseases are the main drivers of shorter lifespan[8] and clear racial differences exist in the prevalence of cardio-metabolic outcomes, such as hypertension, diabetes, metabolic syndrome

and cardiovascular disease[12–19]. However, racial disparities can also be conceptualized from a life-course perspective[20–22] as the differences in the timing, or average age of onset, of a particular outcome. Earlier occurrence of major diseases can lead to earlier disability and death, leading to well-documented and persistent racial/ethnic differences in lifespan[8–11]. The "weathering" hypothesis provides a unifying theory of how earlier onset of disease, rather than overall difference in prevalence, can be caused by the repeated lifetime experience of social/economic adversity and structural marginalization[23]. In support of this hypothesis, various studies have found that cardio-metabolic risk factors and outcomes, such as hypertension, diabetes, stroke, and cardiovascular disease, tend to have a higher prevalence at earlier ages in Blacks compared to Whites[18,23–25]. When racial differences in age of onset of diseases exist, cohorts that restrict eligibility to a particular age range and select individuals free of the outcome (left truncation) or exclude from analyses participants who had the outcome at baseline, and thus at younger ages (left censoring), could result in differential selection by race/ethnicity biasing estimates of racial disparities.

Much effort has focused on mitigating the effect of right censoring[6,26,27], but less research has focused on mitigating potential bias caused by left truncation and left censoring[3–5,7]. Using the Study of Women's Health Across the Nation (SWAN), a multi-ethnic and multi-site longitudinal cohort study of the natural history of menopause, and its cross-sectional screening survey used to recruit participants, this analysis examines the effects of the three forms of selection bias defined above. SWAN selected women into the cohort that were pre- or early peri-menopausal excluding women who reported factors that would preclude observing natural menopause (i.e., use of hormones, being pregnant, having had a hysterectomy and/or bilateral oophorectomy). Cardio-metabolic outcomes were not the main outcome of the study and were not part of the eligibility criteria for selection. Thus, instances of left truncation, left censoring and right censoring for cardio-metabolic outcomes occurred allowing examination of the contribution of each type of selection on estimation of racial/ethnic differences in the timing of cardio-metabolic disease onset. This analysis corrects for each source of selection using a combination of inverse probability weighting (left truncation), models that account for left censoring, and multiple imputation to account for right censoring) and examines how each source of selection affects estimation of the racial/ethnic differences in time to event and predicted median age of onset of 4

cardio-metabolic outcomes including hypertension, isolated systolic hypertension, insulin resistance and diabetes.

4.2 Methods

The Study of Women's Health Across the Nation (SWAN). This analysis uses data from the 1995-1997 cross-sectional screening survey (n = 15,695) and the resulting SWAN cohort study (n = 3,302) with follow up until 2015-2016. The SWAN study is a multi-ethnic multi-site cohort assembled to study women as they aged through the menopausal transition.[28] The cross-sectional screening study recruited women aged 40-55 years that self-identified as the racial/ethnic group for the each site (White at all sites, Black at Detroit, MI; Boston, MA; Chicago, IL; Pittsburgh, PA, Chinese at Oakland, CA, Hispanic at Newark, New Jersey and Japanese at Los Angeles, CA), spoke one of the languages designated for the site (English at all sites plus Cantonese at Oakland, CA, Japanese at Los Angeles, CA, and Spanish at Newark, New Jersey), had a primary residence near the site and could provide informed consent[28,29]. Each screening participant was interviewed regarding reproductive, demographic and health history which determined eligibility for the cohort. The eligibility criteria were: age 42-52 years old, not on hormone therapy for the past 3 months (including birth control, fertility drugs, estrogens or progestins, hormone patches or creams, hormone injections or post-menopausal hormones), not currently pregnant, have an intact uterus and at least 1 ovary (no hysterectomy and/or oophorectomy) and pre-peri menopausal (had a menstrual period in the past 3 months). Participants enrolled in the cohort participated in a baseline and up to 16 follow-up interviews and clinic visits approximately yearly. The Institutional Review Boards at each study site approved the protocol and participants provided informed consent for the screening survey and all cohort visits.

Racial/Ethnic Group. Self-reported primary racial/ethnic group was collected during the screening study. Participants identified themselves as "Black" which included Black, African American or someone of African origin or descent, "Chinese" which included those of Chinese origin or Chinese Americans, "Hispanic" which included Puerto Rican, Mexican/Mexican American, Dominican, Central American, Cuban/Cuban American, South American/Spanish/Other Hispanic, "Japanese" which included those of Japanese origin or Japanese Americans, and "White" which included those identifying as Caucasian.

Hypertension (HTN). A certified technician using standardized protocol took blood pressure (BP) measurements and asked about hypertensive medication use approximately yearly from 1995-1997 (baseline) to 2015-2016 (total of 16 visits). Self-reported anti-hypertensive medication use was verified by medication review at each visit. Participants were asked to refrain from smoking or caffeine 30 minutes before and were seated with feet flat on the floor for 5 minutes before measurement. A standard mercury sphygmomanometer was used to take 3 measurements at least 2 minutes apart. The second and third measurements were averaged to obtain the average systolic and diastolic BP for the session. Stage 1 HTN was defined as systolic BP ≥ 130 mmHg, diastolic BP ≥ 80 mmHg, and/or use of anti-hypertensive medication[30]. Stage 2 HTN was defined as systolic BP ≥ 140 mmHg, diastolic BP ≥ 90 mmHg, and/or use of anti-hypertensive medication[30]. Stage 1 isolated systolic HTN was defined as systolic BP ≥ 130 mmHg and diastolic BP < 80 mmHg or use of anti-hypertensive medication. Stage 2 isolated systolic HTN was defined as systolic BP ≥ 140 mmHg and diastolic BP < 90 mmHg or use of anti-hypertensive medication.

Insulin Resistance. Fasting blood samples were taken and insulin measured using solid-phase radioimmunoassay (DPC Coat-A-Count Insulin RIA; Diagnostic Products, Los Angeles, CA) in 1995-1997 (baseline) and 10 approximately yearly follow up visits (total of 11 visits). The homeostasis model assessment for insulin resistance (HOMA-IR)[31,32] was used to measure insulin resistance; insulin resistance was defined as a HOMA-IR value > 5.9 mlU/L[33,34] or use of insulin medication (self-report verified with medication review).

Diabetes. At each visit from 1995-1997 (baseline) to 2015-2016 (total of 16 visits) type II diabetes status was determined by meeting the one of the following: 10 hour fasting serum glucose level ≥7 mmol/L (available for 11 visits), use of insulin or oral hypoglycemic medication (self-reported verified with medication review) or self-reported physician diagnoses of diabetes.

Covariates. All covariates were measured at baseline. The highest level of education achieved was self-reported (18 categories from did not go to school – doctoral degree) and responses were collapsed into 3 categories (≤ high school education, some college, or an associate degree, ≥ a

college degree). Financial hardship was assessed by asking, “How hard is it for you to pay for the *very basics* like food, housing, medical care, and heating?”, responses were “very hard”, “somewhat hard” and “not very hard at all”. General health was rated on a 5-point scale from poor to excellent: answers were collapsed into a 4-level variable from excellent to fair/poor. Smoking status was reported as: never a smoker, former smoker, and current smoker. Waist circumference was calculated with a measuring tape placed horizontally at the level of the natural waist or the narrowest part of the torso. Body mass index (BMI) was calculated as weight in kilograms divided by measured height in meters squared, measurements were split into categories including less than 25, 25-29.9 and greater than or equal to 30. Self-reported alcohol intake (drinks per week) was collapsed into categories “none to low” (< 2 drinks/week), “moderate” (2-7 drinks/week) and “high” (>7 drinks/week). Physical activity was assessed using a modified Baecke questionnaire and included sport and exercise activity, non-sport leisure activity, and household and childcare activity[35]: scores ranged from 3 to 15.

Multiple Imputation. Multiple imputation by chained equations (MICE) was used to estimate the values of missing data in follow-up[36]. The imputation was performed in two steps. First, all covariates (Fraction of Missing Information (FMI) ≤ 3.5%), use of insulin or hypertension medication (FMI ≤ 5%), occurrence of each outcome at baseline (FMI ≤ 40%) and whether each missing participant had each of the outcomes in follow up (FMI < 40%) were imputed (10 imputation sets). Across all outcomes, approximately < 9% were missing due to interval censoring and < 35% were missing due to loss to follow up (appendix table 1). Predictors in the first imputation included: racial/ethnic group, study site, age at baseline, educational status, financial hardship, smoking status, physical activity score, alcohol use, body mass index, waist circumference, self-reported overall health, employment status, depression score (CESD) and occurrence of everyday discrimination (EDS) with an interaction between racial/ethnic group and educational status, and racial/ethnic group and age. Then, within each of the 10 imputation sets, age of onset of each outcome (age at first occurrence of the outcome) was imputed incorporating the imputed data from the first imputation as predictors. Predictors of the second imputation included: all the predictors in the first imputation plus occurrence of each of the outcomes at baseline and follow up.

Weighting for Left Truncation. For each imputed dataset, three inverse probability weights[2,37,38] were estimated to weight the current cohort to represent the cohort that would have been enrolled if 1) there were no eligibility criteria applied at study screening (eligibility weight), 2) there was no non-response, or differential participation (participation weight) once eligible and 3) everyone in the cohort was recruited at age 42 (study design weight). Data from the cross-sectional screening were used to calculate the eligibility and participation weight, which were calculated as the inverse odds of eligibility and inverse odds of participation, given eligibility. Predictors for each included: race/ethnicity, age, study site, level of education, marital status, HT use, parity, diagnoses of fibroids, self-reported health, diagnosed heart attack/angina, diagnosed osteoporosis, diagnosed cancer (other than skin cancer), smoking status, BMI with 2-way interactions between race/ethnicity and educational level. Age was excluded as a predictor in the eligibility weight to avoid "double counting" the effect of age, as it is incorporated into the study design weight. Predictors in the final models were selected apriori and using forward-backward selection (accuracy for eligibility was 0.80 and participation 0.65). The study design weight was proportional to the 10-year probability of each woman (age 42 to 52 years) being excluded from the study given her age and eligibility status. Weights were multiplied to account for all selection mechanisms simultaneously.

Descriptive Statistics and Accelerated Failure Time Modeling. Mean, proportions and standard errors were calculated and pooled across the imputation sets for all covariates, medication use and for the prevalence and incidence of each outcome in the total sample and across race/ethnic groups.

Accelerated failure time (AFT) models were used to estimate the racial/ethnic differences in age at onset of four cardio-metabolic outcomes (stage 2 HTN, stage 2 isolated systolic HTN, insulin resistance and diabetes). The survival time for participants that had the outcome during observation was equal to age at onset of the condition. For participants that had the outcome at baseline, survival time was left censored. For participants that did not have the outcome during observation time, survival time was right censored. Based on log-likelihood measures the lognormal distribution was the best fit to the data (for all outcomes) thus the interpretation of the regression parameter estimates, with calculation $((\exp(\beta^{coefficient})-1) \times 100)$, are the percent

difference in mean time to each outcome by racial/ethnic group. The racial/ethnic differences in time to each outcome were estimated in four models, where each subsequent model added an additional correction for selection: unadjusted for selection (left censored, right censored, and unweighted for left truncation), incorporating left censored cases (left censored cases considered interval censored and had the outcome between age 20 and age at baseline), accounting for right censoring (uses multiply imputed data to fill in right censored individuals due to interval censoring or loss to follow up), and adjusted for left truncation (weighted for left truncation). All models were adjusted for covariates that were selected a priori, which included educational level, self-reported health, waist circumference, smoking status, alcohol use and physical activity score.

For each AFT model, predicted median age of onset and confidence intervals overall and for each racial/ethnic group were calculated. To account for the effects of covariates, the predicted median age was calculated setting the covariate profile to educational level (some college), self-reported health (very good), waist circumference (mean = 86.34 cm), smoking status (never), alcohol use (moderate) and physical activity score (mean = 7.65). All analyses were conducted in STATA version 16 and R version 4.0.3 (survival and survey packages).

4.3 Results

Table 1 shows multiply imputed descriptive statistics for the SWAN cohort. At baseline, the mean age of participants was 46.3 (SE = 0.05) years old, this is similar across racial/ethnic groups. The plurality of the sample had a college degree or more (42.6%). Black women had higher proportions reporting having some college education (41.4%) and Hispanic women who had a higher proportion reporting having a high school degree or lower (72.1%, $p < 0.05$). At baseline, most women rated their health as "very good" (36.3%) and this was true for all racial/ethnic groups, other than Black and Hispanic women who had higher proportions of women who rated their health as "good" (35.7% and 46.2% respectively, $p < 0.05$). The mean waist circumference at baseline overall was 86.4cm (SE = 0.28), where Japanese women had the lowest mean waist circumference (73.5cm (SE = 0.52)) and Black and Hispanic had the highest (93.1 (SE = 0.54) and 88.2 (SE = 0.83), respectively, $p < 0.05$). Most of the overall sample at baseline had never smoked (58.0%) ranging from 94.4% of Chinese women to 51.7% of White

women ($p < 0.05$). Half of the sample at baseline reported drinking less than 2 alcoholic drinks per week (49.9%); this ranged from 39.6% of White women to 79.1% of Chinese women ($p < 0.05$). The mean physical activity score at baseline (ranges from 3-15 with higher scores meaning more activity) was 7.6 overall (SE = 0.03) and ranged from 6.8 (SE = 0.09) for Hispanic women to 8.0 (SE = 0.05) for White women ($p < 0.05$). Only 4.7% of the sample was using insulin at any time during follow up. White women had the highest proportion reporting insulin use (8.0%) and Japanese women had the lowest proportion (0.5%, $p < 0.05$). Nearly half of the sample overall was using anti-hypertensive medication at any time during follow up (49.4%). White women had the highest proportion reporting anti-hypertensive medication use (67.9%), and Chinese women had the lowest proportion (35.2%, $p <0.05$).

Table 2 shows the proportion of the sample with each of the 4 cardio-metabolic outcomes at baseline and by the end of follow up, overall and by racial/ethnic group. Nearly half (44.5%) of the sample had stage 1 HTN at baseline and 89.6% had it by the end of follow up, compared to 23.9% at baseline and 77.1% by the end of follow up for stage 2 HTN. Black and Hispanic women had the highest proportions of stage 2 HTN at baseline (40.1% and 28.0%, respectively) and by the end of follow up (92.5% and 93.4%, respectively) compared to much smaller proportions for Chinese and Japanese women at baseline (13.2% and 11.7%, respectively) and by the end of follow up (54.6% and 66.2%, respectively). Nearly half of the sample had stage 1 isolated systolic HTN at baseline (44.6%) and end of follow up (79.4%), with only a slight decrease in proportions for stage 2 isolated systolic HTN at baseline (43.7%) and by the end of follow up (72.5%). Black women had the highest proportions of stage 2 isolated systolic HTN at baseline (62.6%) followed by all other racial groups which had around 30% of cases at baseline (range from Hispanic = 26.6% to White = 38.5%). Hispanic women had the largest increase in stage 2 isolated systolic hypertension by the end of follow up (from 26.6% to 88.7%), followed by Black women (from 62.6% to 89.7%). Chinese and Japanese women had the smallest proportions of stage 2 isolated systolic hypertension throughout (from 62.8% to 51.4% and from 36.3% to 57.7%, respectively).

Overall, 13.5% of the sample had insulin resistance at baseline and 35.9% at follow up. Hispanic followed by Black women had the highest proportions of insulin resistance at baseline (28.7%

and 18.7%, respectively) and follow up (64.2% and 51.1% respectively). Japanese and White women had the lowest proportions of insulin resistance at baseline (4.3% and 9.5% respectively) and White women had a higher proportion acquire insulin resistance in follow up (27.5%). Only about 5% of the sample had diabetes at baseline (4.6%) and 24.2% in follow up. Black followed by Hispanic women had the highest proportions of diabetes at baseline (8.6% and 5.9% respectively) and Hispanic women followed by Black women had the highest proportions of diabetes in follow up (45.1% and 34.9% respectively). Japanese and Chinese women had the lowest proportions of diabetes at baseline (0.4% and 1.6% respectively) and follow up (12.1% and 15.9% respectively).

Table 3 displays the percent difference in survival time and the predicted mean age at onset from AFT models, for stage 2 hypertension and isolated systolic hypertension in the four successive models correcting for selection and adjusting for covariates.

For classic HTN, in unadjusted models, Black, Hispanic, and Japanese women develop HTN -9.3% (-12.0, -6.5), -3.1% (-9.5, 3.7) and -2.7% (-6.6, 1.3) earlier while Chinese women develop HTN 3.5% (-1.4, 8.5) later than White women. This corresponds to a predicted median age of onset 5.8 years earlier for Black women, nearly 2 years earlier for Hispanic and Japanese women (1.9 and 1.7 respectively) and 2.1 years later for Chinese women compared to White women. When left censored cases are incorporated, Black women are predicted to have a 14.9% (-17.9, -11.9) and Hispanic women a 16.7% (-22.4, -10.5) earlier onset of HTN compared to White women. This corresponds to a predicted median age of onset that is 8.5 years and 9.5 years earlier for Black and Hispanic women. After accounting for right censoring, Black and Hispanic women have a smaller (-12.1% [-14.1, 10.2] and -10.0% [-13.2, -6.8]) but still significant earlier onset of HTN compared to White women corresponding to a predicted median age of onset 6.9 and 5.7 years earlier than White women, respectively. Similarly, after adjusting for left truncation, Black and Hispanic women have a smaller but still significant earlier onset of HTN (-9.7% [-12.5, -7.0] and -9.8% [-13.8, -5.9]) compared to White women corresponding to a predicted median age of onset 5.4 and 5.5 years earlier than White women. Overall, the median age of onset for HTN decreased by 6.8 years (-7.4, -6.2) after considering all selection mechanisms and Hispanic women had the largest decrease in predicted median age of onset (9.7

years [-11.8, -7.6]). Incorporating left censoring contributed to the largest decrease in predicted median age of onset for all racial/ethnic groups (-6.6 years overall [-6.8, -6.3]) with the largest decreases in Hispanic and Black women (-12.5 years [-13.1, -11.8] and -7.6 years [-7.7, -7.6] respectively). Accounting for right censoring slightly increased the predicted median age of onset overall (+0.8 years [-0.3, 1.8]) especially for Hispanic and Black women (+3.7 years [2.0, 5.5] and +1.5 years [0.5, 2.5] respectively). And accounting for left truncation tended to decrease the predicted median age of onset overall (-1.0 years [-1.3, -0.7]) except for Black women (+0.3 years [0.0, 0.6]).

For isolated systolic HTN, in unadjusted models, all racial/ethnic groups develop isolated systolic HTN earlier than White women, particularly Black women who developed HTN 19.5% (-27.6, -10.5) earlier corresponding to a predicted median age of onset 18.7 years earlier. The difference was reduced after incorporating left censored cases where Black and Hispanic women had -18% (-23.1, 13.5) and -12.6% (-23.2, -0.5) earlier onset than White women corresponding to a predicted median age of onset 8.9 and 6.1 years earlier than White women, respectively. After accounting for right censoring, Black women remained with a -16.0% (-19.4, -12.6) earlier onset and Chinese women had a 5.9% (0.1, 11.6) later onset of HTN compared to White women corresponding to a predicted median age of onset that is 8.5 years earlier for Black women and 3.1 years later for Chinese women. There was little change after adjusting for left truncation, where Black women had a -14.9% (-18.9, -10.9) earlier onset and Chinese women had a 7.3% (0.3, 14.2) later onset of HTN compared to White women corresponding to a predicted median age of onset that is 7.7 years earlier for Black women and 3.8 years later for Chinese women. Overall, the median age of onset for isolated systolic HTN decreased by 42.7 years (-30.9, -54.5) after considering all selection mechanisms, with Black and Hispanic women having the largest decrease in predicted median age of onset (-44.2 years [-32.4, -55.9] and -41.0 years [-15.0, -67.0] respectively). Incorporating left censoring contributed to the largest decrease in predicted median age of onset for all racial/ethnic groups (46.8 years overall [-36.0, -57.6]) with the largest decreases in Hispanic and White women (51.0 years [-27.6, -74.5] and 47.5 years [-36.6, -58.3] respectively). Accounting for right censoring increased the predicted median age of onset, by 5.1 years overall [3.8, 6.4] and especially for Hispanic women (+10.2 years [7.6, 12.6]). Lastly,

accounting for left truncation tended to decrease the predicted median age overall (-0.9 years [-1.2, -0.6]).

For insulin resistance, in unadjusted models, Black and Hispanic women developed insulin resistance -9.1% (-14.6, -3.3) and -18.0% (-27.0, -8.0) earlier than White women, corresponding to a predicted median age of onset 9.5 years and 18.8 years earlier, respectively. After incorporating left censored cases, the earlier onset for Black and Hispanic women increased to -10.1% (-16.4, -3.3) and -27.2% (-35.7, -17.6) earlier than White women, corresponding to a predicted median age of onset 10.8 and 29.1 years earlier. Accounting for right censoring increased the gap between Black and White women to -11.2% (-15.2, -7.2) and slightly decreased the gap between Hispanic and White women to -25.8% (-32.2, -19.4), corresponding to a predicted median age of onset 9.4 and 21.6 years earlier for Black and Hispanic women respectively. Lastly, accounting for left truncation decreased the gap between both groups and White women, where Black and Hispanic women developed insulin resistance -6.2% (-12.2, -0.2) and -25.1% (-34.2, -16.0) earlier than White women corresponding to a predicted median age of onset 5.0 years and 20.1 years earlier. Overall, the median age of onset for HTN decreased by 25.0 years (-27.0, -23.0) after considering all selection mechanisms. Incorporating left censoring lead to a small increase in predicted median age of onset across racial/ethnic groups (0.9 [0.1, 1.8]), with an increase for White and Black women (2.5 [1.2, 3.8] and 1.3 [0.1, 2.4], respectively) and decrease for Hispanic women (-7.7 [-8.0, -7.5]). Accounting for right censoring decreased the predicted median age overall by 23 years (-28.9, -17.1) with the largest decreases for White (-23.2 years [-29.0, -17.5]) and Black (-21.8 years [-27.8, -15.8]) women. Lastly, accounting for left truncation led to an overall decrease the predicted median overall (-2.8 years [-5.8, 0.1]) except for Black women (+0.9 years [-2.8, 4.6]).

For diabetes, in unadjusted models, Black women developed diabetes -5.4% (-10.0, -0.6) earlier than White women corresponding to a predicted median age of onset 5 years earlier for Black women. After incorporating left censored cases, Black and Hispanic women developed diabetes -8.4% (-14.1, -2.3) and -13.2% (-23.2, -1.8) earlier than White women corresponding to a predicted median age 9.0 and 14.1 years earlier. After accounting for right censoring, the gap between Black and White women decreased to -7.3% (-10.9, -3.7) earlier for Black women and

increased to -13.7% (-19.2, -8.1) earlier for Hispanic women, corresponding to a predicted median age of onset 6.8 years and 12.7 years earlier. Lastly, after accounting for left truncation, the gap between Black and White women decreased further to -6.0% (-12.0, 0.1) earlier for Black women and increased to -16.1% (-25.1, -7.1) earlier for Hispanic women, corresponding to a predicted median age of onset 5.4 years and 14.6 years earlier. Overall, the median age of onset for diabetes decreased by 3.8 years (-9.0, 1.3) after considering all selection mechanisms and Hispanic women had the largest decrease in predicted median age of onset (16.3 years [-17.4, -15.2]). Incorporating left censoring increased the predicted median age of onset for all racial/ethnic groups (10.9 years overall [7.9, 13.9]) with the largest increase for White women (13.1 years [9.6, 16.7]). Accounting for right censoring slightly decreased the predicted median age overall by 12.6 years (-17.5, -7.7) with the largest decrease for White women at 14 years (-8.7, -19.5). Accounting for left truncation decreased the predicted median age again by 2.1 years (-9.2, 4.9) overall.

4.4 Discussion

This study provides a comprehensive assessment of the effect of three types of selection bias common to cohort studies on estimated differences in age at onset of hypertension, isolated systolic hypertension, insulin resistance and diabetes across Black, Chinese, Japanese, Hispanic, and White women. We found that addressing selection results in substantial differences in estimated timing of onset of cardio-metabolic risk factors for all racial/ethnic groups. Addressing left truncation, left censoring and right censoring simultaneously lead to an average 20 year decrease in predicted median age of onset across all cardio-metabolic outcomes. Adjusting for left censoring and left truncation had greater effects on outcomes that typically have earlier age at onset in young adulthood like hypertension and isolated systolic hypertension while accounting for right censoring had greater effects on outcomes with later onsets that were more likely to be observed during the study, like insulin resistance and diabetes. Accounting for selection decreased the differences in age at onset of Black and Hispanic women compared to White women. However, in fully adjusted models, Black and Hispanic women still had significantly earlier onset of all outcomes. Hypertension occurred 5 years earlier for Black and Hispanic women, isolated systolic hypertension was particularly prevalent in Black women who had a predicted median age of onset 7.7 years earlier, and metabolic outcomes (insulin resistance

and diabetes) had an average predicted median age 11.3 years earlier than White women. Results provide further longitudinal evidence of weathering or accelerated health declines[18,23] among Black and Hispanic women independent of individual socioeconomic and health behavior factors, warranting further research on the structural causes of such health disparities.

Accounting for right censoring had a larger effect on metabolic outcomes that tended to have later ages of onset than the age of enrollment into the cohort. Based on age specific prevalence estimates, the average age of onset of hypertensive outcomes is earlier (ages 45-64 years) then metabolic outcomes (ages 55-74 years)[18,30,39,40]. The baseline of the SWAN study was age 42-52 years which would allow the study to reliably observe most metabolic outcomes but not hypertensive outcomes, and observation of the latter would be highly dependent on the age of recruitment into the study. Thus, correcting for right censoring via imputation for hypertension likely imputed more non-cases of hypertension than cases, given that most cases occurred earlier and were not right censored. In contrast, for metabolic outcomes, because cases of diabetes and insulin resistance are likely to occur later (during the study) they were probably more likely to be right censored. Thus, the imputation is likely imputing more cases of diabetes and insulin resistance than non-cases, causing the predicted median age of onset in imputed data to decrease. Further, the trend of increased median age of onset for hypertension outcomes with imputation was more exaggerated in racial/ethnic groups that had earlier onset of hypertension, like Black women, and higher amounts of right censoring, like Hispanic women (appendix table 1). On the other hand the decrease in median age of onset for metabolic outcomes was minimized among Hispanic women, suggesting a potential for informative right censoring[41,42] that may bias the estimates of metabolic outcomes for Hispanic women.

Incorporating left censored cases caused decreases in estimated median age for the earlier onset hypertension outcomes and had little effect or increased estimations for observed later onset metabolic outcomes. When left censoring is accounted for in earlier onset outcomes like hypertension[18,30,39,43,44], that have a large amount of left censoring or prevalence at baseline, it led to large decreases in predicted median age of onset as more cases of the outcome were accounted for in the model. However for later onset/observed metabolic outcomes[18,40], that have less left censoring, accounting for left censoring had little effect on the predicted median age of onset. In

a simulation by Cain et al., researchers modeled the effect of accounting for left truncation and left censoring, on estimates of the age of occurrence of a particular outcome when the outcome was also the variable left truncated/left censored[3]. In this simulation when the outcome, which tended to occur earlier than the study start, was progressively left truncated accounting for left censoring alone produced slightly lower estimates of age of onset of the outcome[3]. Although the current analyses models four outcomes that were not the outcomes used to select or left truncate the SWAN study, results show that hypertensive outcomes that tend to occur earlier than age at cohort recruitment do exhibit similar results observed in the Cain et al. article: Adjusting for left censoring alone decreased the age of onset. For later onset metabolic outcomes, the lower frequency of left censoring suggests that they are similar to the "unbiased" estimates in the Cain et al. simulation[3] with little to no effect from adjustment for left censoring.

For the same reasons as left censoring, hypertension outcomes that tend to occur earlier are more affected by adjustment for left truncation than metabolic outcomes that tend to occur during the SWAN study follow up. Given the high levels of left censoring for hypertension, it is likely that levels of left truncation were also high in the SWAN cohort for this outcome. The hypertension findings extend findings of the Cain et al. simulation[3] indicating that, in the presence of potentially high levels of left truncation adjusting for left truncation for an outcome of interest, even if it is not the outcome causing left truncation, can decrease the predicted median age of onset for the outcome of interest. We found this to be true for some but not all racial/ethnic groups. The predicted median age of hypertension for Black and Hispanic women was unaffected or increased slightly with adjustment for left truncation. It could be that adjusting for left truncation based largely on FMP status has less of an effect on onset of outcomes for Black and Hispanic women, who may be more likely to have the outcomes at earlier ages prior to FMP, than White, Chinese and Japanese women[18,30,39,43,44]. Looking at the individual contribution of each weight for left truncation (eligibility, participation, and study design) on estimates of all the outcomes by racial/ethnic group (appendix table 4) supports this hypothesis. The participation or non-response weight moves the age estimates up for Black and Hispanic women, whereas the eligibility and study design weights, that correct for the FMP based eligibility requirements have little to no effect.

The heterogenous effect of left truncation by racial/ethnic group extends conclusions on left truncation by Mayeda et al[7]. Mayeda et al. used simulations to show the effect of survival bias, a type of left truncation caused by mortality, on the association between educational status and late life cognition[7]. The current analyses estimates the association between racial/ethnic group and onset of cardio-metabolic outcomes where left truncation is caused by FMP related selection. Mayeda et al., found that left truncation was causing an underestimate of the association between educational status and late life cognition[7], however we found that left truncation was causing an overestimate of the association between racial/ethnic group and cardio-metabolic health. The results may differ because, as Mayeda et al. explain, the level of bias depends on the magnitude of the association between the variable causing the left truncation (i.e. FMP) and the outcome (i.e. cardio-metabolic health)[7]. However, in this case, we hypothesize that for groups like Black and Hispanic women who are likely having the outcome prior to selection via FMP[18,30,39,43,44] the association is attenuated, instead leading to an overestimate of the racial/ethnic differences.

Black and Hispanic women have a consistently earlier onset of all four cardio-metabolic conditions compared to White women. On average, hypertension occurs 5 years earlier for Black and Hispanic women and the prevalence of isolated systolic hypertension was a particularly high for Black women who had onset of isolated systolic hypertension 7.7 years earlier than White women. Isolated systolic hypertension has been estimated to begin around age of 60 for women in the US[30,44] and is linked to increased rates of cardiovascular and renal diseases[30]. The findings that the predicted median age of onset of isolated systolic hypertension for all groups was around 50 years, and for Black women as early as 41.9 to 44.1 years old, requires immediate research attention as an important target for intervention in early midlife women. Additionally, high incidence of diabetes and especially insulin resistance for Black and Hispanic women, with occurrence 11.3 years earlier in these groups than White women, are important targets for intervention that should start as early as 50 years old for these high-risk groups. The earlier ages of cardio-metabolic disease onset for Black and Hispanic women despite individual socioeconomic and health behavior factors are consistent with the “weathering” hypothesis and warrant further investigation and interventions targeted toward the structural factors causing consistent accelerated differences in health declines.

This study has some limitations. Left truncation may not have been completely accounted for because the inverse probability weights operate under the assumption[2,37] that there are women representative of those who did not make it into the cohort in the retained cohort. This may not fully be the case for SWAN (Aim 2) and the level of representativeness may differ by racial/ethnic group given differential probabilities of eligibility and participation in the cohort (cite Aim 1). There is a large amount of right censoring for Hispanic women compared to other racial/ethnic groups, thus estimates for Hispanic women especially for metabolic outcomes should be interpreted with caution. Lastly, the average prevalence/incidence of metabolic outcomes for Chinese and Japanese women are 10 percent. Thus, the lower predicted median ages of onset for insulin resistance and diabetes compared to White women for these groups should be interpreted with caution and noted that the differences are not statistically significant in the AFT models.

This study also has several strengths. It is the first study to simultaneously correct for three types of selection biases common in cohort studies, examining their varying effects on racial/ethnic differences in onset of cardio-metabolic outcomes. The two-step imputation approach for right censoring is unique in that it allows the predicted age of onset of cardio-metabolic conditions to vary on whether a participant had the outcome during the study period (used as a predictor). We also demonstrate a novel application of inverse probability weighting to correct for left truncation, given that the SWAN had a major strength of collecting information on women who were screened for the cohort[28]. Furthermore, the direct inverse probability weight that weights women in the cohort as if they were all recruited at age 42 (study design weight) is also a novel calculation that can be easily adopted by any cohort of aging to address left truncation stemming from different ages of entry into the cohort.

In conclusion, our results suggest that ignoring selection bias can lead to falsely high estimates of the typical age of onset for these leading causes of morbidity and mortality[8], that have a greater impact on Black and Hispanic women. Not considering the full extent of selection bias in cohort studies can misinform researchers and clinicians on the timing of study and interventions particularly for these high-risk groups that need renewed and continued attention.

Table 4.1 Baseline Characteristics of SWAN, Multiply Imputed Data (n = 3302)

	Total	White (n = 1551)	Black (n = 935)	Chinese (n = 250)	Hispanic (n = 286)	Japanese (n = 281)	p-value[b]
Racial/Ethnic Group[a]	-	47.0	28.3	7.6	8.7	8.5	-
Baseline Characteristics							
Age (Mean, SE)	46.3 (0.05)	46.3 (0.07)	46.2 (0.09)	46.5 (0.16)	46.3 (0.16)	46.7 (0.16)	0.176
Educational Level*							0.000
<= high school	25.1	16.1	26.8	29.0	72.1	18.3	
some college	32.3	30.6	41.4	21.8	18.8	34.3	
>= college	42.6	53.3	31.9	49.2	9.1	47.4	
Financial Hardship							0.000
very hard	9.3	6.0	12.5	5.2	26.4	3.6	
somewhat hard	30.7	26.2	33.8	22.9	55.1	26.7	
not very hard	60.0	67.8	53.6	71.9	18.5	69.8	
Self-Reported Health*							0.000
excellent	21.3	29.2	15.1	16.8	4.9	19.3	
very good	36.3	42.2	32.8	29.4	21.7	36.8	
good	29.2	22.1	35.7	32.4	46.2	26.1	
fair/poor	13.2	6.5	16.4	21.4	27.1	17.8	
Waist Circumference (cm)*	86.4 (0.28)	85.7 (0.41)	93.1 (0.54)	77.3 (0.65)	88.2 (0.83)	73.5 (0.52)	0.000
Smoking Status*							0.000
never	58.0	51.7	52.9	94.4	66.8	68.6	
former	24.7	32.2	22.7	3.4	14.4	20.0	
current	17.3	16.2	24.4	2.2	18.8	11.4	
Body Mass Index							0.000
< 25	39.8	42.8	18.3	76.4	22.8	79.0	
25-29.9	26.2	25.3	28.9	18.4	39.6	15.7	
>= 30	34.0	31.9	52.8	5.2	37.6	5.3	
Alcohol Consumption (servings/week) *							0.000
none/low (< 2)	49.9	39.6	57.0	79.1	50.7	56.1	
moderate (2-7)	28.6	30.6	26.6	14.9	41.6	22.8	
high (>7)	21.6	29.8	16.4	6.0	7.7	21.1	
Physical Activity Score*	7.6 (0.03)	8.0 (0.05)	7.3 (0.06)	7.3 (0.11)	6.8 (0.09)	7.9 (0.10)	0.000
Medication Use through Visit 15							
Insulin	4.7	8.0	4.4	1.0	2.8	0.5	0.000
Hypertension	49.4	67.9	43.1	35.2	44.8	39.5	0.000

Note: [a]Racial/Ethnic group was not imputed. bDifference between racial/ethnic groups. Standard errors between imputation sets were all < 0.05.
* = included as covariate in subsequent models.

Table 4.2 Prevalence and Incidence of Four Cardio-Metabolic Outcomes by Racial/Ethnic Group

	Classic HTN				Isolated Systolic HTN							
	Stage 1[a]		Stage 2[b]		Stage 1[c]		Stage 2[d]		Insulin Resistance		Diabetes	
	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up	Baseline	Follow-Up
Overall	44.5	89.6	23.9	77.1	44.6	79.4	43.7	72.5	13.5	35.9	4.6	24.2
Racial/Ethnic Group												
White	35.7	86.6	17.3	70.4	39.5	73.3	38.5	65.3	9.5	27.5	3.2	17.6
Black	56.5	97.2	40.1	92.5	64.3	93.5	62.6	89.7	18.7	51.1	8.6	34.8
Chinese	28.4	73.9	13.2	54.6	32.8	61.6	32.8	51.4	11.7	18.4	1.6	15.9
Hispanic	78.3	100.0	28.0	93.4	26.6	93.9	26.6	88.7	28.7	64.2	5.9	45.1
Japanese	33.8	84.7	11.7	66.2	36.3	67.2	36.3	57.7	4.3	18.2	0.4	12.1

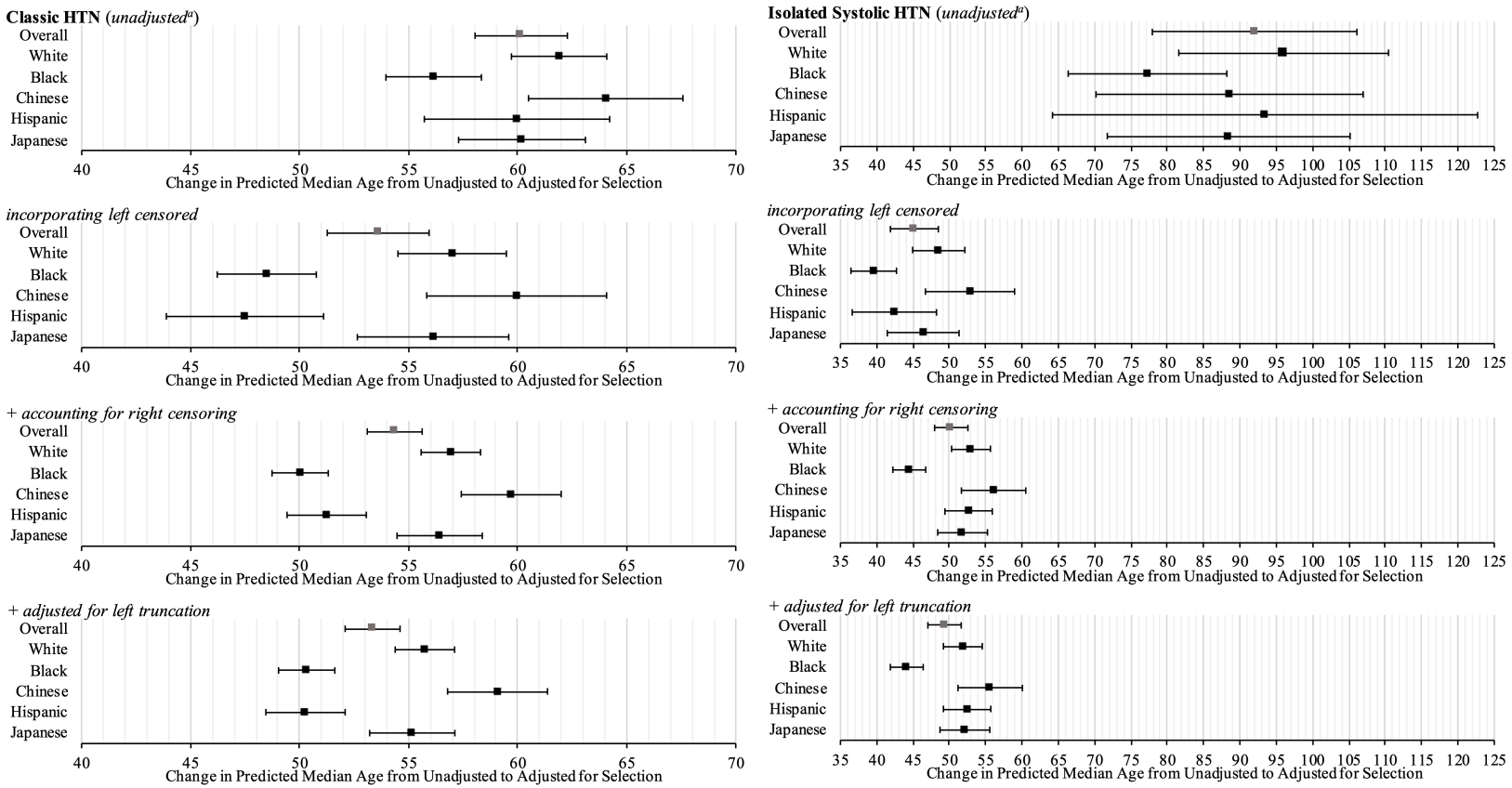
Note: HTN = hypertension. Standard errors between imputation sets were all < 0.05.
[a]Classic HTN (Stage 1) = Systolic blood pressure >= 130, Diastolic blood pressure >= 80 and/or use of HTN medication.
[b]Classic HTN (Stage 2) = Systolic blood pressure >= 140, Diastolic blood pressure >= 90 and/or use of HTN medication.
[c]Isolated Systolic HTN (Stage 1) = Systolic blood pressure >= 130 and Diastolic blood pressure < 80 or use of HTN medication.
[d]Isolated Systolic HTN (Stage 2) = Systolic blood pressure >= 140 and Diastolic blood pressure < 90 or use of HTN medication.

Table 4.3 Percent Difference in Survival Time and Predicted Median Age at Onset of Hypertension (Stage 2) by Racial/Ethnic Group

	Classic HTN							Isolated Systolic HTN						
	%	LCI	UCI	p-value	Median	LCI	UCI	%	LCI	UCI	p-value	Median	LCI	UCI
Unadjusted[a]														
Overall					60.2	58.0	62.3					92.0	77.9	106.1
Racial/Ethnic Group														
White					61.9	59.7	64.1					96.0	81.6	110.4
Black	-9.3	-12.0	-6.5	0.001	56.1	54.0	58.3	-19.5	-27.6	-10.5	0.006	77.3	66.3	88.2
Chinese[b]	3.5	-1.4	8.5	0.101	64.0	60.5	67.6	-7.8	-21.7	8.7	0.177	88.6	70.2	106.9
Hispanic	-3.1	-9.5	3.7	0.188	60.0	55.7	64.2	-2.7	-27.1	29.9	0.426	93.4	64.2	122.7
Japanese[b]	-2.7	-6.6	1.3	0.108	60.2	57.3	63.1	-7.9	-20.8	7.1	0.155	88.4	71.7	105.1
Incorporating Left Censored														
Overall					53.6	51.3	55.9					45.2	41.9	48.5
Racial/Ethnic Group														
White					57.0	54.5	59.5					48.5	44.9	52.1
Black	-14.9	-17.9	-11.9	0.000	48.5	46.2	50.8	-18.4	-23.1	-13.5	0.001	39.6	36.4	42.7
Chinese[b]	5.2	-1.0	11.8	0.075	60.0	55.8	64.1	8.9	-1.7	20.5	0.075	52.8	46.7	59.0
Hispanic	-16.7	-22.4	-10.5	0.003	47.5	43.9	51.1	-12.6	-23.2	-0.5	0.045	42.4	36.6	48.2
Japanese[b]	-1.5	-6.6	3.8	0.284	56.1	52.6	59.6	-4.4	-12.8	4.7	0.176	46.4	41.4	51.3
+ Accounting for Right Censoring														
Overall					54.4	53.1	55.6					50.3	48.3	52.2
Racial/Ethnic Group														
White					56.9	55.6	58.3					53.0	50.8	55.2
Black	-12.1	-14.1	-10.2	0.000	50.0	48.7	51.3	-16.0	-19.4	-12.6	0.000	44.5	42.5	46.5
Chinese[b]	4.9	1.5	8.2	0.003	59.7	57.4	62.0	5.9	0.1	11.6	0.026	56.1	52.4	59.8
Hispanic	-10.0	-13.2	-6.8	0.000	51.2	49.4	53.1	-0.6	-5.9	4.6	0.408	52.6	49.5	55.8
Japanese[b]	-0.9	-3.9	2.1	0.278	56.4	54.5	58.4	-2.2	-7.4	3.0	0.204	51.8	48.7	54.9
+ Adjusted for Left Truncation														
Overall					53.3	51.8	54.9					49.3	47.0	51.6
Racial/Ethnic Group														
White					55.7	53.9	57.6					51.8	49.2	54.5
Black	-9.7	-12.5	-7.0	0.000	50.3	48.7	52.0	-14.9	-18.9	-10.9	0.000	44.1	41.9	46.4
Chinese[b]	6.0	1.8	10.2	0.003	59.1	56.1	62.1	7.3	0.3	14.2	0.024	55.6	51.2	60.0
Hispanic	-9.8	-13.8	-5.9	0.000	50.3	48.1	52.4	1.1	-4.1	6.4	0.337	52.4	49.2	55.7
Japanese[b]	-1.0	-6.4	4.4	0.352	55.2	51.7	58.6	0.5	-5.1	6.2	0.426	52.1	48.7	55.6

Note: Percent difference calculated as (exp(β)-1)*100.
All models adjusted for covariates including educational level, self-reported health, waist circumference, smoking status, alcohol use and physical activity score.
Covariates for predicted median age were set to educational level (some college), self-reported health (very good), waist circumference (86.34 cm), smoking status (never), alcohol use (moderate) and physical activity score (7.65).
[a]Data in model is left censored, right censored and unadjusted for left truncation.
[b]Small number of cases for Chinese (n with insulin resistance ~ 46, n with diabetes ~ 40) and Japanese (n with insulin resistance ~ 51, n with diabetes ~ 34) women.

Figure 4.1 Predicted Median Age of Onset of Hypertension (Stage 2) Overall and by Racial/Ethnic Group



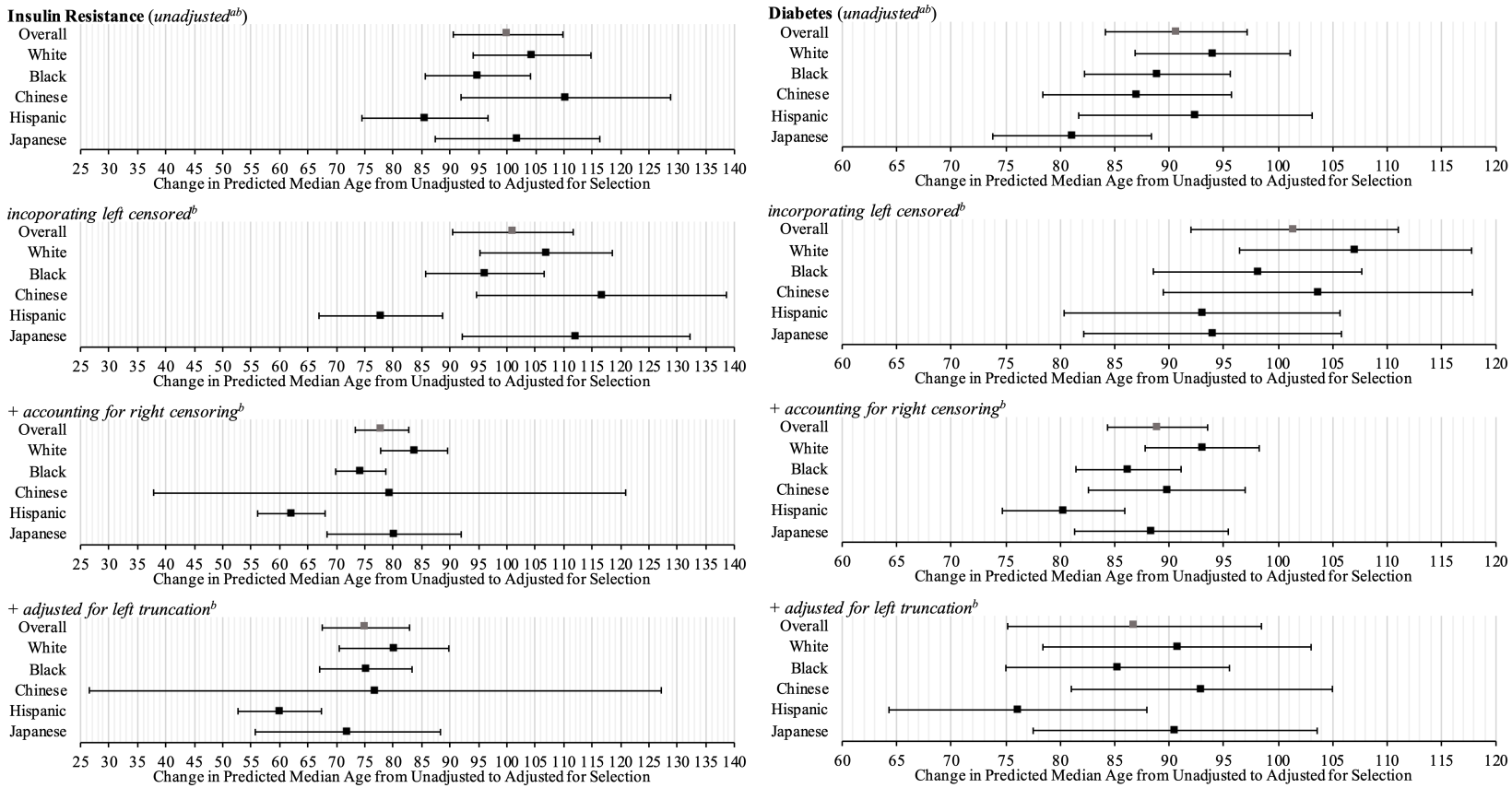
Note: HTN = hypertension. Covariates for predicted median age were set to: educational level (some college), self-reported health (very good), waist circumference (86.34 cm), smoking status (never), alcohol use (moderate) and physical activity score (7.65).
[a]Data in model is left censored, right censored, and unadjusted for left truncation.

Table 4.4 Percent Difference in Survival Time and Predicted Median Age at Onset of Metabolic Outcomes by Racial/Ethnic Group

	Insulin Resistance							Diabetes						
	%	LCI	UCI	p-value	Median	LCI	UCI	%	LCI	UCI	p-value	Median	LCI	UCI
Unadjusted[a]														
Overall					100.2	90.5	109.8					90.6	84.1	97.1
Racial/Ethnic Group														
White					104.4	94.0	114.8					94.0	86.9	101.1
Black	-9.1	-14.6	-3.3	0.016	94.9	85.6	104.1	-5.4	-10.0	-0.6	0.038	88.9	82.2	95.6
Chinese[b]	5.7	-8.6	22.1	0.230	110.3	91.9	128.8	-7.4	-14.9	0.8	0.063	87.1	78.4	95.7
Hispanic	-18.0	-27.0	-8.0	0.011	85.6	74.5	96.6	-1.7	-11.4	9.2	0.374	92.4	81.7	103.1
Japanese[b]	-2.4	-13.3	9.8	0.339	101.8	87.4	116.3	-13.7	-20.1	-6.8	0.007	81.1	73.8	88.4
Incorporating Left Censored														
Overall					101.1	90.5	111.7					101.5	92.0	111.0
Racial/Ethnic Group														
White					106.9	95.3	118.5					107.1	96.5	117.8
Black	-10.1	-16.4	-3.3	0.018	96.1	85.7	106.5	-8.4	-14.1	-2.3	0.022	98.1	88.6	107.7
Chinese[b]	9.1	-7.7	28.9	0.164	116.6	94.7	138.6	-3.2	-13.8	8.6	0.288	103.6	89.5	117.8
Hispanic	-27.2	-35.7	-17.6	0.003	77.8	66.9	88.7	-13.2	-23.2	-1.8	0.035	93.0	80.4	105.7
Japanese[b]	4.9	-10.2	22.6	0.273	112.2	92.1	132.2	-12.2	-21.2	-2.3	0.030	94.0	82.2	105.8
+ Accounting for Right Censoring														
Overall					78.0	73.3	82.7					88.9	84.4	93.5
Racial/Ethnic Group														
White					83.7	77.8	89.6					93.0	87.8	98.3
Black	-11.2	-15.2	-7.2	0.000	74.3	69.9	78.7	-7.3	-10.9	-3.7	0.000	86.3	81.5	91.1
Chinese[b]	-5.2	-17.2	6.7	0.180	79.4	37.9	120.9	-3.5	-10.3	3.3	0.152	89.8	82.6	97.0
Hispanic	-25.8	-32.2	-19.4	0.000	62.1	56.2	68.0	-13.7	-19.2	-8.1	0.000	80.3	74.7	85.9
Japanese[b]	-4.2	-12.7	4.3	0.158	80.2	68.4	92.0	-5.0	-11.7	1.7	0.066	88.4	81.3	95.4
+ Adjusted for Left Truncation														
Overall					75.2	67.5	82.9					86.8	75.2	98.5
Racial/Ethnic Group														
White					80.1	70.5	89.8					90.7	78.4	103.0
Black	-6.2	-12.2	-0.2	0.019	75.2	67.1	83.3	-6.0	-12.1	0.1	0.024	85.3	75.0	95.5
Chinese[b]	-4.2	-17.9	9.5	0.262	76.8	26.5	127.2	2.5	-6.3	11.3	0.290	93.0	81.0	105.0
Hispanic	-25.1	-34.2	-16.0	0.000	60.0	52.7	67.4	-16.1	-25.1	-7.1	0.000	76.1	64.3	88.0
Japanese[b]	-10.1	-27.4	7.1	0.112	72.0	55.7	88.3	-0.2	-9.0	8.6	0.483	90.6	77.5	103.6

Note: Percent difference calculated as (exp(β)-1)*100.
All models adjusted for covariates including educational level, self-reported health, waist circumference, smoking status, alcohol use and physical activity score.
Covariates for predicted median age were set to educational level (some college), self-reported health (very good), waist circumference (86.34 cm), smoking status (never), alcohol use (moderate) and physical activity score (7.65).
[a]Data in model is left censored, right censored, and unadjusted for left truncation.
[b]Small number of cases for Chinese (n with insulin resistance ~ 46, n with diabetes ~ 40) and Japanese (n with insulin resistance ~ 51, n with diabetes ~ 34) women.

Figure 4.2 Predicted Median Age of Onset of Metabolic Conditions by Racial/Ethnic Group



Note: HTN = hypertension. Covariates for predicted median age were set to: educational level (some college), self-reported health (very good), waist circumference (86.34 cm), smoking status (never), alcohol use (moderate) and physical activity score (7.65).
[a]Data in model is left censored, right censored, and unadjusted for left truncation.
[b]Small number of cases for Chinese (n with insulin resistance ~ 46, n with diabetes ~ 40) and Japanese (n with insulin resistance ~ 51, n with diabetes ~ 34) women.

Table 4.5 Missing Patterns across Four Cardio-Metabolic Outcomes through Visit 15 [appendix]

	Observed[a]	Missing Interval	Loss to Follow Up
	% (n)	% (n)	% (n)
Insulin Resistance			
Total	61.5 (2030)	8.5 (280)	30.1 (993)
Racial/Ethnic Group			
White	64.9 (1006)	5.7 (88)	29.5 (457)
Black	63.6 (595)	13.3 (124)	23.1 (216)
Chinese	54.0 (135)	7.6 (19)	38.4 (96)
Hispanic	47.6 (136)	11.2 (32)	41.3 (118)
Japanese	56.2 (158)	6.1 (17)	37.7 (106)
Diabetes			
Total	66.2 (2186)	2.0 (65)	31.9 (1052)
Racial/Ethnic Group			
White	67.9 (1053)	1.2 (19)	30.9 (479)
Black	66.2 (619)	2.0 (19)	31.8 (297)
Chinese	78.0 (195)	0.0 (0)	22.0 (55)
Hispanic	39.2 (112)	9.4 (27)	51.4 (147)
Japanese	73.7 (207)	0.0 (0)	26.3 (74)
Classic HTN			
Total	67.9 (2241)	8.5 (280)	23.7 (782)
Racial/Ethnic Group			
White	66.0 (1024)	8.1 (126)	25.9 (401)
Black	76.3 (713)	8.5 (79)	15.3 (143)
Chinese	70.8 (177)	4.8 (12)	24.4 (61)
Hispanic	45.5 (130)	17.5 (50)	37.1 (106)
Japanese	70.1 (197)	4.6 (13)	25.3 (71)
Isolated Systolic HTN			
Total	64.6 (2132)	8.9 (294)	26.6 (877)
Racial/Ethnic Group			
White	64.2 (996)	7.5 (117)	28.2 (438)
Black	72.3 (676)	9.6 (90)	18.1 (169)
Chinese	70.4 (176)	4.4 (11)	25.2 (63)
Hispanic	33.6 (96)	21.7 (62)	44.8 (128)
Japanese	66.9 (188)	5.0 (14)	28.1 (79)

Note: HTN = hypertension.
[a]Observed fully or carried forward (due to no change in status through interval).

Table 4.6 Percent Difference in Survival Time and Predicted Median Age at Onset of Hypertension by Racial/Ethnic Group Unadjusted for Covariates by Left Truncation Weight [appendix]

	Classic							Isolated Systolic						
	%	LCI	UCI	p-value	Median	LCI	UCI	%	LCI	UCI	p-value	Median	LCI	UCI
All Selection Accounted for Other than Left Truncation														
Overall			-		54.2	53.7	54.7					50.8	50.1	51.6
Racial/Ethnic Group														
White			*ref*		57.3	56.6	58.0					54.4	53.2	55.5
Black	-16.8	-18.9	-14.8	0.000	47.7	46.9	48.4	-23.2	-26.6	-19.7	0.000	41.8	40.6	42.9
Chinese[a]	8.3	5.0	11.6	0.000	62.1	60.1	64.0	12.1	6.5	17.8	0.000	61.0	57.7	64.2
Hispanic	-13.9	-17.0	-10.9	0.000	49.3	47.9	50.7	-7.0	-12.0	-1.9	0.002	50.6	48.3	52.9
Japanese[a]	4.2	1.2	7.3	0.004	59.7	58.1	61.4	6.4	1.1	11.8	0.011	57.9	55.0	60.7
Adjusted for Eligibility Only														
Overall					53.6	53.1	54.1					49.5	48.7	50.3
Racial/Ethnic Group														
White					56.9	56.0	57.7					53.2	51.9	54.5
Black	-16.3	-18.6	-14.0	0.000	47.6	46.8	48.4	-16.6	-23.8	-9.3	0.000	41.3	40.2	42.3
Chinese[a]	8.5	4.8	12.3	0.000	61.7	59.5	63.9	9.9	1.2	18.6	0.016	60.3	56.7	64.0
Hispanic	-13.0	-16.1	-9.9	0.000	49.5	48.2	50.8	-25.6	-38.8	-12.5	0.000	50.4	48.5	52.2
Japanese[a]	5.1	1.9	8.3	0.001	59.8	58.0	61.5	18.0	8.2	27.8	0.000	57.5	54.2	60.8
Adjusted for Participation Only														
Overall					54.0	53.5	54.4					51.0	50.2	51.7
Racial/Ethnic Group														
White					56.9	56.1	57.7					54.3	53.1	55.5
Black	-15.4	-17.6	-13.3	0.000	48.1	47.4	48.9	-22.4	-25.9	-18.9	0.000	42.1	41.1	43.2
Chinese[a]	9.4	5.7	13.0	0.000	62.2	60.1	64.4	12.5	6.2	18.7	0.000	61.1	57.5	64.7
Hispanic	-12.7	-15.7	-9.8	0.000	49.7	48.4	50.9	-5.6	-9.9	-1.3	0.004	51.3	49.4	53.2
Japanese[a]	4.9	1.8	8.0	0.001	59.7	58.0	61.4	6.3	0.6	12.1	0.018	57.8	54.7	60.8
Adjusted for Study Design Only														
Overall					53.1	52.3	53.8					49.8	48.8	50.8
Racial/Ethnic Group														
White					56.4	55.1	57.7					53.4	51.7	55.1
Black	-15.4	-18.8	-12.1	0.000	47.7	46.6	48.7	-22.0	-26.9	-17.1	0.000	41.6	40.1	43.1
Chinese[a]	8.7	4.2	13.2	0.000	61.3	58.7	63.8	11.9	4.4	19.3	0.002	59.7	55.6	63.8
Hispanic	-15.4	-20.0	-10.8	0.000	47.7	45.8	49.5	-6.4	-10.8	-2.0	0.002	50.0	48.5	51.5
Japanese[a]	3.1	-3.0	9.2	0.164	58.1	54.8	61.4	8.3	2.6	14.0	0.003	57.8	55.0	60.6
Fully Adjusted for Left Truncation														
Overall					52.6	52.0	53.2					49.2	48.3	50.1
Racial/Ethnic Group														
White					55.8	54.7	56.9					52.6	51.1	54.1
Black	-14.0	-16.9	-11.2	0.000	48.0	47.1	48.8	-21.0	-25.4	-16.7	0.000	41.6	40.2	42.9
Chinese[a]	9.7	5.5	13.8	0.000	61.2	58.8	63.5	13.2	6.3	20.0	0.000	59.6	55.7	63.4
Hispanic	-13.0	-16.7	-9.2	0.000	48.6	47.1	50.1	-4.0	-8.7	0.7	0.045	50.5	48.6	52.4
Japanese[a]	4.4	-1.2	10.0	0.067	58.2	55.1	61.3	9.2	3.6	14.8	0.001	57.4	54.6	60.2

Note: Percent difference calculated as (exp(β)-1)*100.

[a]Small number of cases for Chinese (n with insulin resistance ~ 46, n with diabetes ~ 40) and Japanese (n with insulin resistance ~ 51, n with diabetes ~ 34) women.

Table 4.7 Percent Difference in Survival Time and Predicted Median Age at Onset of Metabolic Outcomes by Racial/Ethnic Group Unadjusted for Covariates by Left Truncation Weight [appendix]

	Insulin Resistance							Diabetes						
	%	LCI	UCI	p-value	Median	LCI	UCI	%	LCI	UCI	p-value	Median	LCI	UCI
All Selection Accounted for Other than Left Truncation														
Overall					75.9	73.9	77.9					88.3	85.6	91.1
Racial/Ethnic Group														
White					83.4	80.4	86.4					94.7	91.0	98.4
Black	-22.9	-27.4	-18.4	0.000	64.3	62.1	66.5	-19.1	-23.3	-14.9	0.000	76.6	74.0	79.3
Chinese[a]	2.4	-10.9	15.8	0.355	85.5	36.6	134.5	3.9	-3.6	11.4	0.160	98.4	90.8	106.0
Hispanic	-33.9	-40.5	-27.3	0.000	55.1	51.7	58.5	-23.6	-29.6	-17.5	0.000	72.4	68.3	76.4
Japanese[a]	14.0	4.8	23.2	0.003	95.1	77.5	112.7	11.5	3.8	19.2	0.003	105.6	96.8	114.4
Adjusted for Eligibility Only														
Overall					74.2	70.8	77.7					87.1	79.2	95.1
Racial/Ethnic Group														
White					81.6	75.7	87.6					94.0	81.8	106.2
Black	-21.6	-27.9	-15.3	0.000	64.0	61.4	66.6	-18.8	-26.5	-11.0	0.000	76.3	70.7	82.0
Chinese[a]	4.8	-8.3	17.9	0.235	85.6	40.4	130.9	6.2	-2.0	14.4	0.076	99.8	86.5	113.1
Hispanic	-31.8	-40.3	-23.3	0.000	55.7	52.2	59.2	-22.5	-32.2	-12.9	0.000	72.8	67.4	78.2
Japanese[a]	18.9	9.7	28.2	0.000	97.1	78.3	115.9	14.4	6.4	22.4	0.000	107.5	91.8	123.2
Adjusted for Participation Only														
Overall					75.8	72.3	79.2					87.2	81.3	93.1
Racial/Ethnic Group														
White					83.2	77.9	88.5					93.8	86.4	101.1
Black	-20.7	-26.2	-15.2	0.000	66.0	63.3	68.7	-18.2	-23.1	-13.2	0.000	76.8	72.4	81.1
Chinese[a]	3.7	-10.9	18.4	0.305	86.4	26.5	146.3	3.9	-3.7	11.5	0.162	97.4	87.9	107.0
Hispanic	-31.2	-38.9	-23.6	0.000	57.2	53.5	60.9	-22.3	-28.9	-15.8	0.000	72.8	68.1	77.6
Japanese[a]	15.7	6.2	25.3	0.002	96.3	76.0	116.6	12.4	4.8	20.0	0.001	105.4	94.2	116.7
Adjusted for Study Design Only														
Overall					73.0	67.3	78.7					85.6	72.8	98.4
Racial/Ethnic Group														
White					80.5	71.3	89.7					91.0	81.3	100.8
Black	-19.3	-28.6	-10.1	0.000	64.9	60.7	69.2	-16.6	-23.8	-9.3	0.000	75.9	68.6	83.3
Chinese[a]	3.8	-10.5	18.2	0.300	83.7	37.9	129.5	9.9	1.2	18.6	0.016	100.1	87.5	112.6
Hispanic	-36.2	-50.5	-21.9	0.000	51.3	45.4	57.2	-25.6	-38.8	-12.5	0.000	67.7	55.9	79.5
Japanese[a]	5.7	-12.9	24.3	0.279	85.1	64.5	105.7	18.0	8.2	27.8	0.000	107.4	91.2	123.7
Fully Adjusted for Left Truncation														
Overall					72.9	68.2	77.6					84.7	74.1	95.2
Racial/Ethnic Group														
White					80.3	72.3	88.4					90.9	77.8	104.0
Black	-17.3	-25.4	-9.2	0.000	66.4	62.8	70.0	-15.8	-24.1	-7.4	0.000	76.6	69.7	83.5
Chinese[a]	5.8	-8.6	20.2	0.216	85.1	36.2	134.0	11.3	2.7	19.9	0.007	101.1	86.3	116.0
Hispanic	-31.6	-43.2	-19.9	0.000	55.0	49.7	60.3	-22.6	-34.3	-10.9	0.000	70.3	61.2	79.5
Japanese[a]	10.1	-8.6	28.7	0.156	88.4	67.1	109.7	20.3	11.0	29.7	0.000	109.4	90.4	128.4

Note: Percent difference calculated as (exp(β)-1)*100.

[a]Small number of cases for Chinese (n with insulin resistance ~ 46, n with diabetes ~ 40) and Japanese (n with insulin resistance ~ 51, n with diabetes ~ 34) women.

4.5 References

1. Rothman, K. J., Greenland, S. & Lash, T. J. *Modern Epidemiology*. (Lippincott Williams and Wilkins, 2014).
2. Hernan, M. A. & Robins, J. M. *Causal Inference*. (2016).
3. Cain, K. C. *et al.* Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes. *Am. J. Epidemiol.* **173**, 1078–1084 (2011).
4. Schisterman, E. F., Cole, S. R., Ye, A. & Platt, R. W. Accuracy loss due to selection bias in cohort studies with left truncation. *Paediatr. Perinat. Epidemiol.* **27**, 491–502 (2013).
5. Howards, P. P., Hertz-Picciotto, I. & Poole, C. Conditions for bias from differential left truncation. *Am. J. Epidemiol.* **165**, 444–452 (2007).
6. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A structural approach to selection bias. *Epidemiology* **15**, 615–625 (2004).
7. Mayeda, E. R., Filshtein, T. J., Tripodis, Y., Maria Glymour, M. & Gross, A. L. Does selective survival before study enrolment attenuate estimated effects of education on rate of cognitive decline in older adults? A simulation approach for quantifying survival bias in life course epidemiology. *Int. J. Epidemiol.* **47**, 1507–1517 (2018).
8. Kochanek, K. D., Arias, E. & Anderson, R. N. *How Did Cause of Death Contribute to Racial Differences in Life Expectancy in the United States in 2010? NCHS data brief, no 125* (2013).
9. Nan, |, Astone, M., Ensminger, M. & Juon, H. S. *Early Adult Characteristics and Mortality Among Inner-City African-American Women*. *American Journal of Public Health* **92**, (2002).
10. Geronimus, A. T., Bound, J., Waidmann, T. A., Hillemeier, M. M. & Burns, P. B. Excess Mortality among Blacks and Whites in the United States. *N. Engl. J. Med.* **335**, 1552–1558 (1996).
11. Geronimus, A. T., Bound, J., Waidmann, T. A., Colen, C. G. & Steffick, D. Inequality in Life Expectancy, Funcational Status, and Active Life Expectancy Across Selected Black and White Populations in the United States. *Demography* **38**, 227–251 (2001).
12. Singh, G. K., Siahpush, M., Azuine, R. E. & Williams, S. D. Widening Socioeconomic and Racial Disparities in Cardiovascular Disease Mortality in the United States ,. **3**, 106–118 (2015).
13. Bell, C. N., Thorpe, R. J., Bowie, J. V. & LaVeist, T. A. Race disparities in cardiovascular disease risk factors within socioeconomic status strata. *Ann. Epidemiol.* **28**, 147–152 (2018).
14. Fuller-Rowell, T. E. *et al.* Disparities in insulin resistance between black and white adults in the United States: The role of lifespan stress exposure. *Psychoneuroendocrinology* **107**, 1–8 (2019).
15. Hasson, B. R., Apovian, C. & Istfan, N. Racial/Ethnic Differences in Insulin Resistance and Beta Cell Function: Relationship to Racial Disparities in Type 2 Diabetes among African Americans versus Caucasians. *Curr. Obes. Rep.* **4**, 241–249 (2015).
16. Lin, S. X., Carnethon, M., Szklo, M. & Bertoni, A. Racial/ethnic differences in the association of triglycerides with other metabolic syndrome components: The multi-ethnic study of atherosclerosis. *Metab. Syndr. Relat. Disord.* **9**, 35–40 (2011).
17. Falkner, B. & Cossrow, N. D. F. H. Prevalence of metabolic syndrome and obesity-associated hypertension in the racial ethnic minorities of the United States. *Curr.*

Hypertens. Rep. **16**, (2014).
18. Thorpe, R. J. *et al.* Accelerated Health Declines among African Americans in the USA. *J. Urban Heal.* **93**, 808–819 (2016).
19. Geronimus, A. T., Bound, J., Keene, D. & Hicken, M. Black-White Differences in Age Trajectories of Hypertension Prevalence Among Adult Women and Men, 1999-2002. *Ethn. Dis.* **17**, 40–48 (2007).
20. Gee, G. C., Hing, A., Mohammed, S., Tabor, D. C. & Williams, D. R. Racism and the Life Course: Taking Time Seriously. *Am J Public Heal.* **109**, 43–47 (2019).
21. Ben-Shlomo, Y. & Kuh, D. A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *Int. J. Epidemiol.* **31**, 285–293 (2002).
22. Braveman, P. What is health equity: And how does a life-course approach take us further toward it? *Matern. Child Health J.* **18**, 366–372 (2014).
23. Geronimus, A. T., Hicken, M., Keene, D. & Bound, J. 'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am. J. Public Health* **96**, 826–833 (2006).
24. Levine, M. E. & Crimmins, E. M. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc. Sci. Med.* **118**, 27–32 (2014).
25. Geronimus, A. T. *et al.* Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-based Sample. *J. Health Soc. Behav.* **56**, 199–224 (2015).
26. Mayeda, E. R. *et al.* A Simulation Platform for Quantifying Survival Bias: An Application to Research on Determinants of Cognitive Decline. *Am. J. Epidemiol.* **184**, 378–387 (2016).
27. Howe, C. J., Cole, S. R., Lau, B., Napravnik, S. & Eron, J. J. Selection bias due to loss to follow up in cohort studies. *Epidemiology* **176**, 139–148 (2017).
28. Sowers, M. *et al.* SWAN: A Multicenter, Multiethnic, Community Based Cohort Study of Women and the Menopausal Transition. in *Menopause: Biology and Pathobiology* 175–188 (2000).
29. Gold, E. B. *et al.* Factors Associated with Age at Natural Menopause in a Multiethnic Sample of Midlife Women. *Am. J. Epidemiol.* **153**, 865–874 (2001).
30. Chobanian, A. V *et al.* Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatement of High Blood Pressure. *Hypertension* **42**, 1206–1252 (2003).
31. Hanson, R. L. *et al.* Evaluation of simple indices of insulin sensitivity and insulin secretion for use in epidemiologic studies. *Am. J. Epidemiol.* **151**, 190–198 (2000).
32. Bonara, E. *et al.* Homeostasis Model Assessment Closely Mirrors the Glucose Clamp Technique in the Assessment of Insulin Sensitivity. *Diabetes Care* **23**, 57 (2000).
33. Gayoso-Diz, P. *et al.* Insulin resistance (HOMA-IR) cut-off values and the metabolic syndrome in a general adult population: effect of gender and age: EPIRCE cross-sectional study. *BMC Endocr. Disord.* **13**, 47 (2013).
34. Bardini, G., Barbaro, V., Romano, D., Rotella, C. M. & Giannini, S. Different distribution of phenotypes and glucose tolerance categories associated with two alternative proposed cutoffs of insulin resistance. *Acta Diabetol.* **51**, 321–324 (2014).
35. Baecke, J. Determinants of Body Fatness in Young Adults Living in a Dutch Community. *Doctoral Dissertation Baecke* (1982).
36. Rubin, D. B. Multiple Imputations in Sample Surveys - A Phenomenological Baysian

Approach to Nonresponse. *Proc. Surv. Res. methods Sect. Am. Stat. Assoc. Am. Stat. Assoc.* **1**, 20–34 (1978).
37. Seaman, S. R. & White, I. R. Review of inverse probability weighting for dealing with missing data. doi:10.1177/0962280210395740
38. Seaman, S. R., White, I. R., Copas, A. J. & Li, L. Combining Multiple Imputation and Inverse-Probability Weighting. *Biometrics* (2012). doi:10.1111/j.1541-0420.2011.01666.x
39. Burt, V. L. *et al.* Prevalence of Hypertension in the US Adult Population. *Hypertension* **25**, 305–313 (1995).
40. Brancati, F. L., Kao, W. H. L., Folsom, A. R., Watson, R. L. & Szklo, M. Incident Type 2 Diabetes Mellitus in African American and White Adults. *JAMA* **283**, 2253 (2000).
41. Moore, D. F. *Applied Survival Analysis Using R*. (Springer, 2016). doi:10.1007/978-3-319-31245-3
42. Siannis, F., Copas, J. & Lu, G. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics* **6**, 77–91 (2005).
43. Zeki, A. *et al.* Use of a pooled cohort to impute cardiovascular disease risk factors across the adult life course. *Int. J. Epidemiol.* **0**, (2018).
44. Kannel, W. B. *et al.* Systolic Blood Pressure, Arterial Rigidity, and Risk of Stroke. *J. Am. Med. Assoc.* **245**, 1225 (1981).

Chapter 5 Conclusions

An emerging body of research suggests that Black and Hispanic women "weather" or experience accelerated health declines compared to White women in the US, potentially explaining persistent racial differences in lifespan[1–6]. However, this information has not been adequately incorporated into the formation of many cohorts of aging[7–10], that are assembled with the underlying assumption that "normative" aging is what occurs in White populations. Most cohorts of aging recruit people at ages before the "typical" age of onset of the outcome of interest as they aim to observe the occurrence of the outcome during the study[11,12]. However, as some racial groups are experiencing many adverse health conditions earlier than the White population, on which the typical age of onset is often based[13], a systematic bias or selection bias may arise.

Selection bias is a distortion in estimation caused by differential selection into or out of a study or analyses[11,14] If selection into or out of a cohort study affects racial/ethnic groups differentially due to "weathering"[3], selection would preclude correct estimation of the risk of health outcomes being observed with differential errors across race/ethnic groups. Therefore, the extent of racial disparities would also be measured with error impacting the study of potential mitigating factors. Although a growing body of research addresses multiple forms of selection bias in cohorts of aging via theory and simulations[12,15–17], little research has empirically assessed multiple sources of selection bias and examined the differential impact of selection on racial/ethnic groups.

This dissertation uses novel data from the screening study and cohort of the Study of Women's Health Across the Nation (SWAN), a multi-ethnic 25-year longitudinal cohort in the US aimed at observing the menopausal transition[18], to fill this critical research gap. Three aims were undertaken to fully examine the impact of selection bias on estimation of racial/ethnic disparities in the SWAN cohort. The first aim created a framework for examining selection bias in a cohort

and quantified the extent that selection may bias racial/ethnic differences in SWAN. The second aim corrected for all forms of selection bias and estimated their effects on racial/ethnic differences on a main outcome of the SWAN study, the timing of the final menstrual period (FMP). The third aim corrected for all forms of selection bias and estimated their effects on racial/ethnic differences in the timing of various secondary cardio-metabolic outcomes in the SWAN study, including hypertension, isolated systolic hypertension, insulin resistance and diabetes.

All aims used the cross-sectional screening survey, designed to recruit the SWAN cohort as well as data from the SWAN cohort. The community-based screening study included 15,695 women who were 40-55 years old, self-identified as the racial/ethnic group for the designated site, spoke either English or the language for the designed site, had a primary residence in the designated geographic area for each site and were able to give informed consent. The SWAN cohort enrolled 3202 women from the study screening that were ages 42-52 years old, not on hormone therapy, not currently pregnant, had an intact uterus and at least 1 ovary, and were pre-peri menopausal. Data were multiply imputed and pooled across 10 imputation sets to ensure no missing variables for cohort participants.

5.1 Summary of Findings

Aim 1. As a starting point to mitigating selection bias in the SWAN cohort chapter two provided an applied framework for approaching selection bias in a cohort of aging. First, SWAN procedures were reviewed to elucidate the mechanisms of selection at play in the cohort. Two main mechanisms were identified, eligibility and participation. The framework focused particularly on left truncation i.e., when individuals who have already had the outcome of interest are not included in the study, as a main cause of bias given that accelerated health declines are likely to affect inclusion into the cohort. We then quantify the extent to which selection may affect estimation of racial/ethnic disparities in the timing of menopause and related outcomes in SWAN. Second, the association between various potential predictors, including racial/ethnic group, of each stage of selection were estimated. Examining if health related factors were highly associated with inclusion and if inclusion varied by racial/ethnic group, we effectively estimate the potential for left truncation to bias racial/ethnic disparities in the cohort.

Two main selection mechanisms were identified. The eligibility mechanism (1) selects women from the screening study who met the eligibility criteria for the cohort. Within the eligibility mechanism is nested an inherent feature of the study design which limits eligibility to women that happen to be recruited at the right combination of age (between 42-52 years) and eligibility status. The participation mechanism (2) selects eligible women from the cross-sectional screening who are also willing and able to participate into the cohort.

Eligibility differed greatly by racial/ethnic group. Black and Hispanic women had the lowest eligibility rates (38.8% and 40.5%). Left truncation mostly stemmed from high rates of surgical menopause (hysterectomy and/or oophorectomy) (30.9% and 17.9% respectively). Demographic and health related predictors such as age, fibroid diagnoses, body mass index, smoking status and previous diagnoses of diabetes, heart attack/angina, osteoporosis and cancer were strongly associated with eligibility while demographic predictors alone, such as educational level and financial hardship, were strongly associated with participation. Thus, we observed no evidence to suggest a "healthy volunteer bias" caused by differential participation rates, but the eligibility selection mechanism, which varied greatly by racial/ethnic group, may have had a considerable influence on selection bias and left truncation.

A strong interaction existed between racial/ethnic group and age in predicting eligibility. Increases in age were associated with larger decreases in the odds of eligibility for Black and Hispanic women compared to White, Chinese, and Japanese women. This shows that among Black and Hispanic women who had particularly high levels of left truncation, left truncation increased at a higher rate with age than in other racial/ethnic groups providing evidence of "weathering" or accelerated health declines[2,3].

The proposed directed acyclic graph (DAG) for the effect of selection on the association between racial/ethnic group and FMP (or other related outcomes) shows that selection via FMP status creates collider stratification bias, or bias in estimation from stratifying on a variable (selection in this case) that is affected by two or more variables related to the exposure/outcome process[15,19]. A strong interaction between racial/ethnic group and education was present also.

Among White women, decreasing levels of education were associated with lower odds of eligibility but no education trend was apparent for other racial/ethnic groups. Previous research has shown that strong interactions between the exposure (racial/ethnic group) and variables predicting selection cause more bias than when no interaction is present[15]. Given the proposed causal structure and evidence of interaction, these results show that bias in racial/ethnic differences in SWAN are likely considerable.

Overall, chapter 2 exemplifies a framework for approaching selection at study commencement in a cohort of aging and presents strong evidence that current estimates that do not account for selection at study commencement or for left truncation may underestimate racial/ethnic disparities in menopausal timing and other related outcomes.

Aim 2. Chapter three examined the effect of selection by correcting for the relevant selection biases and estimating effects of these corrections on estimates of the racial/ethnic differences in final menstrual period (FMP) timing, the main outcome for the SWAN study. FMP status was the variable used to determine presence of left truncation, as individuals who had already had their FMP were not included in the study[18]. Previous SWAN research had shown no racial/ethnic differences in FMP timing (with adjustment)[20], thus the goal of this aim was to re-estimate the racial/ethnic differences in FMP timing accounting for all selection biases including left truncation.

Two relevant sources of selection biases in the estimation of FMP were corrected for in this analysis. Left truncation caused by eligibility and participation, identified in Chapter 2, were addressed using inverse probability weighting. Essentially inverse probability weights weight the cohort as if no eligibility criteria had been applied and no non-participation had occurred. Right censoring was the second form of selection bias addressed, caused by loss to follow up or missing data on FMP status in the retained cohort.

Right censoring was addressed using multiple imputation[21]. Additionally, any missing-ness on the covariates of interest which included sociodemographic and health behavior factors were also multiply imputed. Multiple imputation was undertaken in two steps, first to determine FMP type

(surgical or natural) and then to determine FMP timing along with the other covariates with FMP type as a predictor. FMP age was imputed within the likely times of occurrence for each participant using bounded imputation[22,23]. Left truncation was the second source of selection addressed. This was done with a combination of three inverse probability weights. Two weights utilized the SWAN cross-sectional study screening and 1) upweighted women in the retained cohort representative of women who were not eligible (eligibility weight) and 2) upweighted eligible women in the retained cohort representative of women who were unwilling or unable to participate (participation weight). The last weight utilized information from within the cohort, upweighting women in the retained cohort as if they were all recruited into the study at age 42, instead of between the ages 42-52 years. The three weights were multiplied to fully account for left truncation. Cox proportional hazard models, with age as the timescale, were used to estimate the racial/ethnic differences in timing to FMP, with moderation by type of FMP (natural or surgical).

Results showed that previous estimates in SWAN that ignored the effects of selection bias via left truncation and right censoring offer falsely high estimates of the average age of FMP and thus underestimate racial/ethnic differences in FMP timing. Selection biases particularly affected Black women whose age of natural FMP was overestimated by an average of 1.1 years prior to accounting for all forms of selection bias. Overall once selection biases were adjusted for, Black women had natural menopause 0.60 years earlier and surgical menopause 1.8 years earlier than White women and disparities for Hispanic women, although not significant, followed similar trends.

To the best of our knowledge, this was one of the first studies to simultaneously correct for left truncation and right censoring in a cohort of aging. Although mitigation is imperfect, results show that not considering selection biases, particularly stemming from left truncation, biases estimation of the timing of FMP with larger impact for Black and Hispanic women. Results also show consistent evidence of earlier natural and surgical FMP among Black and Hispanic women compared to White women that is present independent of socioeconomic factors and health behaviors. This chapter concludes that racial differences may be driven by a combination of "weathering"[3] and overuse of reproductive surgeries in Black and Hispanic women.

Aim 3. The last step to examining the effect of selection bias on racial/ethnic disparities was to correct for all relevant selection biases and estimate their effect on racial/ethnic differences in the timing of secondary cardio-metabolic outcomes in the SWAN study. The menopausal transition ending with the occurrence of the FMP, is a time when women are posited to be at highest cardio-metabolic risk due to loss of cardio-protective and metabolic regulating reproductive hormones[24–28]. Because cardio-metabolic outcomes are not primary outcomes of the SWAN study, a third type of selection is introduced. Left censoring occurs when individuals who have had the outcome of interest are included in the study[12] and thus have the outcome of interest, in this case cardio-metabolic outcomes, at baseline. Like left truncation and right censoring, left censoring can cause bias if the probability of left censoring is differential between racial/ethnic groups. The goal of this aim was to correct for all three forms of selection bias present when estimating secondary outcomes, including right censoring, left truncation and left censoring, examining their effects on racial/ethnic differences in the age of onset of hypertension, isolated systolic hypertension, insulin resistance and diabetes.

Blood pressure measurements, fasting venous blood draws and medication review were made approximately yearly in the SWAN study for approximately 20 years. Hypertension (HTN) was defined as systolic BP ≥ 140 mmHg, diastolic BP ≥ 90 mmHg, and/or use of anti-hypertensive medication[29]. Isolated systolic HTN was defined as systolic BP ≥ 140 mmHg and diastolic BP < 90 mmHg or use of anti-hypertensive medication. Insulin resistance was measured via the homeostasis model assessment for insulin resistance (HOMA-IR)[30,31] and defined as a HOMA-IR value > 5.9 mlU/L[32,33] or use of insulin medication (self-report verified with medication review). Diabetes was defined as 10 hour fasting serum glucose level ≥7 mmol/L, use of insulin or oral hypoglycemic medication (self-reported verified with medication review) or self-reported physician diagnoses of diabetes.

Right censoring was accounted for by using multiple imputation to fill in missing outcomes or covariates due to loss to follow up. The imputation was performed in two steps. First, all covariates and whether each of the outcomes occurred throughout the study were imputed. Second, the age of occurrence of the outcome was imputed using the covariate values including

those variables imputed in the first imputation as predictors. Left truncation was accounted for in the same way as in Aim 2, with inverse probability weights that weighted the cohort as if it was not left truncated. Left censoring was accounted for in Accelerated Failure Time (AFT) models that can account for left censoring. The resulting AFT models are interpreted as the percent difference in the mean time to each outcome by racial/ethnic group (with a lognormal distribution).

Results showed that accounting for all forms of selection bias led to an average 20 year decrease in predicted median age of onset across all cardio-metabolic outcomes overall. Adjusting for left censoring and left truncation had greater effects for earlier onset outcomes like hypertension and isolated systolic hypertension, whereas right censoring had greater effects on later onset outcomes that were observed during the study, like insulin resistance and diabetes. Accounting for selection decreased the predicted degree of earlier onset for Black and Hispanic women compared to White women. However, in fully adjusted models, Black and Hispanic women still had significant earlier onset of all outcomes. Hypertension occurred 5 years earlier for Black and Hispanic women, isolated systolic hypertension was particularly prevalent in Black women who had a predicted median age of onset 7.7 years earlier, and metabolic outcomes (insulin resistance and diabetes) had an average predicted median age 11.3 years earlier than White women.

In summary, chapter three found that ignoring selection biases can lead to falsely higher estimations of the age of onset for secondary outcomes overall, however the level of bias was similar across racial/ethnic groups. This may be because most of the selection bias was stemming from left truncation, and the level of bias depends on the magnitude of the association between the variable causing the left truncation (i.e. FMP) and the outcome (i.e. cardio-metabolic health)[7]. However, in this case, we hypothesize that for groups like Black and Hispanic women who likely have the outcome prior to selection via FMP[18,30,39,43,44] the association is attenuated, causing the Black and Hispanic estimates to be less affected by adjustment for left truncation. Regardless Black and Hispanic women had an earlier predicted median onset of all outcomes than White women that was still present after accounting for socioeconomic and health behavior factors. This finding warrants further investigation and suggests the need for targeted early interventions addressing the structural factors causing consistent accelerated health declines.

5.2 Limitations and Strengths

Limitations include the fact that SWAN did not have information on how the target sample population at each site differed from the cross-sectional screening participants. Therefore, generalizability to the target population and/or lack of additional bias stemming from selection into the cross-sectional study cannot be assumed to be negligible. The SWAN study recruited Hispanic, Chinese, and Japanese women at one site each, thus geographic/site differences cannot be disentangled from racial/ethnic differences for these groups. The prevalence of surgical FMP, insulin resistance and diabetes outcomes for Chinese and Japanese women were quite low, thus models in these ethnicities lack adequate power and have imprecise estimates for these groups. The inverse probability weights for left truncation effectively upweight women retained in the cohort who are similar to women who were left truncated; thus, we assume that there are women in the cohort that are representative of left truncated women. However, this assumption may not be completely met in the SWAN cohort given that we observed clear differences in sociodemographic and health characteristics between the screening study and retained cohort (Aim 1). Given this, we conclude that the left truncation weights may not fully account for left truncation caused by initial ineligibility or non-participation (eligibility and participation weights). Lastly, the multiple imputation performed for right censoring relies on the assumption that right censoring is occurring at random. However, Hispanic women have much higher rates of right censoring compared to other racial/ethnic groups thus estimates for this group may be imprecise.

Nonetheless, strengths of this study are notable. SWAN is one of the hallmark studies of women's health through the midlife[10] and contains extensive reproductive, sociodemographic, and detailed health outcomes for up to 25 years of follow up for multiple racial/ethnic groups with excellent retention[18]. In addition, SWAN presented a unique opportunity to examine the extent of left truncation, given that it collected extensive information on women that were screened for entry into the cohort in the cross-sectional screening study. Most cohorts do not have information about selection along with extensive measurements on participants that were screened for the cohort, which makes these aims possible.

Although inverse probability weighting is a common approach to account for confounding, selection bias stemming from right censoring, and to increase the representativeness of a study sample[14,34], to the best of our knowledge, weights have not been used previously to account for bias stemming from left truncation and selection at study commencement. Most analyses leave the issue of left truncation to a short acknowledgement in the limitations section of their manuscripts despite a growing body of simulation research reinforcing the likelihood that biases are clearly instilled when selection at study commencement is ignored[12,15].The novel application of inverse probability weighting to the problem of left truncation is a major strength of this dissertation. Furthermore, the study design inverse probability weight uses information from the cohort rather than depending on representation in the cohort of those with lower odds of inclusion. Although it does not address left-truncation completely, results show that weighting women in the cohort as if they were recruited at age 42 years with the study design weight contributes the most to addressing bias stemming from left truncation. This approach is a novel application of weighting and a major strength, in that the approach could be easily adopted by any cohort study to account for wide ranges of age-eligible for a given study. Lastly, Chapters two and four employed a unique two step multiple imputation procedure that incorporated surgical menopause and occurrence of the outcomes as moderators in predicting age at onset of FMP and the four cardio-metabolic outcomes, respectively. In Chapter two, this was a major strength in that those lost to follow up were not assumed to have natural menopause and FMP age was allowed to vary by FMP type improving the estimates for each. In Chapter 3, this was a major strength as it allowed the predicted age at onset of the outcome to vary by whether the participant had the condition overall improving the accuracy of imputation.

5.3 Public Health Implications

This dissertation exemplifies that the problem of selection bias warrants more attention than to be relegated to a sentence in the limitations section. It provides an applied framework and statistical strategies for dealing with selection biases in a large cohort of aging. It also demonstrates the salience of selection bias for racial/ethnic disparities. Results show that prior estimates that do not consider selection bias underestimate the estimated age of onset of health outcomes, especially for Black and Hispanic women thus underestimating racial/ethnic disparities. Selection bias affecting racial/ethnic differences is an issue that most researchers will

have to grapple with to obtain unbiased estimates of their associations of interest. The framework and methods asserted in this work have implications for the analysis of current cohort data and for newly starting cohorts.

For current cohorts, addressing selection should start with clearly mapping out the causal question of interest i.e., the causal relationship between exposure and outcome identifying important confounders and mediators in a DAG. With the exposure, outcome and important confounding factors clearly defined, the stages of selection for the cohort of interest can be reviewed for potential differential selection. Selection can then be added to the DAG to determine if selection affects the ability to identify the causal association between the exposure and outcome of interest. If in review, the eligibility criteria for the cohort are dependent on being free of the outcome of interest attention should be paid to how eligibility of different groups of interest may be affected, linking the differential association back to the causal question of interest. This dissertation demonstrates eligibility criteria can select against Black and Hispanic racial groups that are experiencing accelerated health declines. Most causal questions, whether they need to control for, or are explicitly looking at racial/ethnic group differences, will need to grapple with differential selection via racial/ethnic group as it may affect the ability to identify a causal association. For SWAN investigators, the main predictors of selection have been determined in chapter 2 which can be easily mapped onto new questions of interest and analytical plans to determine if the association requires a formal accounting for selection.

After mapping out the potential effects of selection to the causal question of interest. Chapter 3 and 4 demonstrate that statistical tools and strategies exist for accounting for selection. Multiple imputation is a common approach used for right censoring. When there is left censoring, left censored cases can be incorporated into flexible survival models via interval censoring. Left truncation, can be largely addressed by inverse probability weighting. In the absence of a screening cohort, left truncation caused by entry into the study at varying ages within the age range for the study (i.e., weighting as if everyone entered at the lowest age of cohort entry) can be accounted for using the direct weighting approach shown in Chapters 2 and 3.

These results also have implications for newly starting cohorts. Although much attention has rightly been focused on non-participation bias, researchers should also be thoughtful about how eligibility criteria may be differentially affecting the groups they intend to generalize to. In particular, the age range that a study is recruiting combined with an eligibility criterion that is dependent on age can inadvertently leave the highest risk groups out of the study of interest. As shown in these chapters, Black and Hispanic women would need to be recruited at much younger ages than other groups to observe the full spectrum of reproductive and cardio-metabolic aging and their inter-relationship in these populations. Furthermore, results show that it is important to collect ample data on potential participants that did not make it into the cohort, if not done already. This data can help quantify the extent of potential bias and mitigate bias that may be present despite thoughtful study planning.

Lastly, these results demonstrate the stark and persistent racial/ethnic disparities and accelerated health declines that span reproductive aging and cardio-metabolic health. The racial disparities reported in SWAN are persistent regardless of socioeconomic standing or health behaviors signaling that individual level interventions will at most only partially mitigate these disparities. In the case of reproductive aging, clear unexplained differences in surgical menopause persist that are consistent with an over-diagnoses and overuse of reproductive surgeries for Black and Hispanic women[36]. Research and interventions, starting well before midlife, targeting structural and individual causes of these disparities are urgently needed.

5.4 Future Research Directions

Three major future research directions emerge from this dissertation's findings.

The Role of the FMP in Cardio-Metabolic Outcomes. Chapter 4 examines the effect of selection bias on estimates of racial/ethnic differences in the timing of cardio-metabolic risk. SWAN selected women based on their FMP status. Thus, correcting for left truncation indirectly examined the effect of FMP on cardio-metabolic timing. However, further research is needed on the explicit effect of FMP timing on cardio-metabolic risk i.e., the association between FMP timing and cardio-metabolic risk. A subsequent question is whether racial/ethnic differences in the timing of cardio-metabolic risk are mediated by FMP timing. And lastly, given large

racial/ethnic differences in the prevalence of surgical FMP, examining the role of natural versus surgical FMP in promoting or mitigating cardio-metabolic risk onset overall and by racial/ethnic group.

Structural Racism and Racial Differences in Accelerated Health Declines. Chapters 3 and 4 found that Black and Hispanic women experienced earlier FMP and accelerated cardio-metabolic health declines despite controling for individual sociodemographic and health behavior factors, consistent with the theory of "weathering". Thus, further work should be aimed at identifying structural racism and race-related stressors as drivers of early reproductive aging and onset of cardio-metabolic risk.

Racial segregation has been characterized as a fundamental structural cause for racial/ethnic disparities in health[3,6,35–44] and has been consistently linked to poor health outcomes and excess mortality. At the individual level, racial discrimination is an increasingly well-studied example of a personally-mediated race-related event shown to be a salient chronic stressor linked to poor health outcomes[3,6,35–44]. Researchers have argued that both institutional and personally-mediated levels of racial discrimination are present at the same time and work to reinforce each other in producing poor health outcomes among racial/ethnic minorities[37,39,42,45]. However, little evidence has been produced to show the pathways by which structural racism, enacted via racial segregation, and personally mediated racism may interact to produce health disparities[46–52]. Future work can utilize the rich data collected in SWAN to further understand the reinforcing pathways by which racial segregation and personal mediated racism interact to cause accelerated health declines for Black and Hispanic women. Furthermore, because Black women are recruited at more than one site and given evidence of differential rates of surgical menopause by geographic location[53], future research should aim to disentangle the effect of geographic location and racial segregation on rates of surgical menopause for Blacks and Whites in SWAN.

Correction for Selection by Cohort Linking. Although much of the bias stemming from left truncation may be mitigated by the approach and methods in this dissertation, there still may be bias present stemming from age ranges before the cross-sectional study took place (before the age of 42 years). This may preclude the SWAN cohort from gauging the full spectrum of the

midlife transition particularly for Black and Hispanic women. A newly developing area of research uses imputation and weighting methods[54] to link existing cohorts spanning different age ranges (ex. cohort 1 contains women ages 20-40 and cohort 2 contains women ages 40-75), so that left truncation can be fully mitigated and a fuller picture of the life-course can be observed, utilizing the strengths and measurements of each cohort. Future work should identify cohorts that would provide similar information on earlier life stages than SWAN and use causal inference driven linking procedures[54] to observe the entire midlife transition for all racial/ethnic groups.

5.5 Conclusion

In conclusion, an understanding of, not only the differences in prevalence, but the *timing of onset* of health declines is essential for focusing research and interventions to the right life stages which may differ by race/ethnicity. This dissertation fills a critical research gap, showing how selection bias in a large cohort of aging leads to erroneous estimations of racial/ethnic differences in the timing of health declines, providing a way forward for understanding and mitigating selection bias in a cohort of aging. Results also show accelerated health declines in Black and Hispanic women across multiple reproductive and cardio-metabolic outcomes, urgently calling for continued and renewed research and interventions aimed at the structural causes of consistent racial disparities in health.

5.6 References

1. Geronimus, A. T., Bound, J., Keene, D. & Hicken, M. Black-White Differences in Age Trajectories of Hypertension Prevalence Among Adult Women and Men, 1999-2002. *Ethn. Dis.* **17**, 40–48 (2007).
2. Thorpe, R. J. *et al.* Accelerated Health Declines among African Americans in the USA. *J. Urban Heal.* **93**, 808–819 (2016).
3. Geronimus, A. T., Hicken, M., Keene, D. & Bound, J. 'Weathering' and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am. J. Public Health* **96**, 826–833 (2006).
4. Forrester, S. *et al.* Racial differences in weathering and its associations with psychosocial stress: The CARDIA study. *SSM - Popul. Heal.* **7**, 100319 (2019).
5. Levine, M. E. & Crimmins, E. M. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc. Sci. Med.* **118**, 27–32 (2014).
6. Geronimus, A. T. *et al.* Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-based Sample. *J. Health Soc. Behav.* **56**, 199–224 (2015).
7. Howard, V. J. *et al.* The reasons for geographic and racial differences in stroke study: Objectives and design. *Neuroepidemiology* **25**, 135–143 (2005).
8. Hill, C. *et al.* The Atherosclerosis Risk in Communities (ARIC) Study: Design and Objectives. **129**, (1989).
9. Juster, F. T. & Suzman, R. An Overview of the Health and Retirement Study. *J. Hum. Resour.* **30**, S7 (1995).
10. Ferrell, R. J. & Sowers, M. Longitudinal, epidemiologic studies of female reproductive aging. *Ann. N. Y. Acad. Sci.* **1204**, 188–197 (2010).
11. Rothman, K. J., Greenland, S. & Lash, T. J. *Modern Epidemiology*. (Lippincott Williams and Wilkins, 2014).
12. Cain, K. C. *et al.* Bias Due to Left Truncation and Left Censoring in Longitudinal Studies of Developmental and Disease Processes. *Am. J. Epidemiol.* **173**, 1078–1084 (2011).
13. Doll, R. Cohort studies: History of the method I. Prospective cohort studies. *Soz. Praventivmed.* **46**, 75–86 (2001).
14. Hernan, M. A. & Robins, J. M. *Causal Inference*. (2016).
15. Mayeda, E. R., Filshtein, T. J., Tripodis, Y., Maria Glymour, M. & Gross, A. L. Does selective survival before study enrolment attenuate estimated effects of education on rate of cognitive decline in older adults? A simulation approach for quantifying survival bias in life course epidemiology. *Int. J. Epidemiol.* **47**, 1507–1517 (2018).
16. Howe, C. J. & Robinson, W. R. Survival-related selection bias in studies of racial health disparities. *Epidemiology* **29**, 521–524 (2018).
17. Howe, C. J., Cole, S. R., Lau, B., Napravnik, S. & Eron, J. J. Selection bias due to loss to follow up in cohort studies. *Epidemiology* **176**, 139–148 (2017).
18. Sowers, M. *et al.* SWAN: A Multicenter, Multiethnic, Community Based Cohort Study of Women and the Menopausal Transition. in *Menopause: Biology and Pathobiology* 175–188 (2000).
19. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A structural approach to selection bias. *Epidemiology* **15**, 615–625 (2004).
20. Gold, E. B. *et al.* Factors related to age at natural menopause: Longitudinal analyses from SWAN. *Am. J. Epidemiol.* (2013). doi:10.1093/aje/kws421

21. Rubin, D. B. Multiple Imputations in Sample Surveys - A Phenomenological Baysian Approach to Nonresponse. *Proc. Surv. Res. methods Sect. Am. Stat. Assoc. Am. Stat. Assoc.* **1**, 20–34 (1978).
22. Kwon, T. Y. & Park, Y. A new multiple imputation method for bounded missing values. *Stat. Probab. Lett.* **107**, 204–209 (2015).
23. Geraci, M. & McLain, A. Multiple Imputation for Bounded Variables. *Psychometrika* **83**, 919–940 (2018).
24. Snowdon, D. A. *et al.* Is Early Natural Menopause a Biologic Marker of Health and Aging? *Am. J. Public Health* **79**, 709–714 (1989).
25. Wise, P. M., Krajnak, K. M. & Kashon, M. L. Menopause: The Aging of Multiple Pacemakers. *Am. Assoc. Adv. Sci.* **273**, 67–70 (1996).
26. Cooper, G. S. & Sandler, D. P. Age at Natural Menopause and Mortality. *Ann. Epidemiol.* **8**, 229–235 (1998).
27. Gold, E. B. The timing of the age at which natural menopause occurs. *Obstet. Gynecol. Clin. North Am.* **38**, 425–40 (2011).
28. Mendelsohn, M. E. & Karas, R. H. The Protective Effects of Estrogen on the Cardiovascular System. *N. Engl. J. Med.* **340**, 1801–1811 (1999).
29. Chobanian, A. V *et al.* Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatement of High Blood Pressure. *Hypertension* **42**, 1206–1252 (2003).
30. Hanson, R. L. *et al.* Evaluation of simple indices of insulin sensitivity and insulin secretion for use in epidemiologic studies. *Am. J. Epidemiol.* **151**, 190–198 (2000).
31. Bonara, E. *et al.* Homeostasis Model Assessment Closely Mirrors the Glucose Clamp Technique in the Assessment of Insulin Sensitivity. *Diabetes Care* **23**, 57 (2000).
32. Gayoso-Diz, P. *et al.* Insulin resistance (HOMA-IR) cut-off values and the metabolic syndrome in a general adult population: effect of gender and age: EPIRCE cross-sectional study. *BMC Endocr. Disord.* **13**, 47 (2013).
33. Bardini, G., Barbaro, V., Romano, D., Rotella, C. M. & Giannini, S. Different distribution of phenotypes and glucose tolerance categories associated with two alternative proposed cutoffs of insulin resistance. *Acta Diabetol.* **51**, 321–324 (2014).
34. Seaman, S. R. & White, I. R. Review of inverse probability weighting for dealing with missing data. doi:10.1177/0962280210395740
35. Lewis, T. T., Williams, D. R., Tamene, M. & Clark, C. R. Self-Reported Experiences of Discrimination and Cardiovascular Disease. *Curr. Cardiovasc. Risk Rep.* **8**, 365 (2014).
36. Williams, D. R. *et al.* Research on discrimination and health: An exploratory study of unresolved conceptual and measurement issues. *Am. J. Public Health* **102**, 975–978 (2012).
37. Williams, D. R. & Mohammed, S. A. Discrimination and racial disparities in health: evidence and needed research. *J. Behav. Med.* **32**, 20–47 (2009).
38. Diet, A., In, D., Hollingswolth, S., Am, A.- & Candelaria, J. G. Levels of Racism and a Gardener's Tale. (2010). doi:10.1016/0014-2999(90)90458-I
39. Paradies, Y. *et al.* Racism as a Determinant of Health: A Systematic Review and Meta-Analysis. *PLoS One* **10**, e0138511 (2015).
40. Berger, M. & Sarnyai, Z. 'More than skin deep': Stress neurobiology and mental health consequences of racial discrimination. *Stress* **18**, (2015).
41. Busse, D., Yim, I. S., Campos, B. & Marshburn, C. K. Discrimination and the HPA axis:

current evidence and future directions. *Journal of Behavioral Medicine* **40**, (2017).
42. Nazroo, J. Y. The structuring of ethnic inequalities in health: Economic position, racial discrimination, and racism. *American Journal of Public Health* (2003). doi:10.2105/AJPH.93.2.277
43. Albert, M. A. & Williams, D. R. Invited commentary: Discrimination - An emerging target for reducing risk of cardiovascular disease? *Am. J. Epidemiol.* **173**, 1240–1243 (2011).
44. Clark, R., Anderson, N., Clark, V. & Williams, D. Racism as a stressor for african american: A Biopsychosocial Model. *Am. Psychol.* (1999).
45. Jones, C. P. *et al.* Using "socially assigned race" to probe white advantages in health status. *Ethn. Dis.* **18**, 496–504 (2008).
46. English, D., Lambert, S. F., Evans, M. K. & Zonderman, A. B. Neighborhood Racial Composition, Racial Discrimination, and Depressive Symptoms in African Americans. *Am. J. Community Psychol.* **54**, 219–228 (2014).
47. Hurd, N. M., Sellers, R. M., Cogburn, C. D., Butler-Barnes, S. T. & Zimmerman, M. A. Racial Identity and Depressive Symptoms among Black Emerging Adults: The Moderating Effects of Neighborhood Racial Composition. *Dev Psychol* **49**, 938–950 (2013).
48. Root, E. D. & Humphrey, J. L. Neighborhood racial composition and trajectories of child self-rated health: An application of longitudinal propensity scores. *Soc. Sci. Med.* **120**, 31–39 (2014).
49. Hunt, M. O., Wise, L., Jipguep, M.-C., Cozier, Y. C. & Rosenberg, L. Neighborhood Racial Composition and Perceptions of Racial Discrimination: Evidence From the Black Women's Health Study*. *Soc. Psychol. Q.* **70**, 272–289 (2007).
50. Russell, E. F. *et al.* Metropolitan area racial residential segregation, neighborhood racial composition, and breast cancer mortality. doi:10.1007/s10552-012-0029-4
51. Gee, G. C. A Multilevel Analysis of the Relationship Between Institutional and Individual Racial Discrimination and Health Status. (2008).
52. Koole, S. L. & Rothermund, K. 'I feel better but I don't know why': The psychology of implicit emotion regulation. (2011). doi:10.1080/02699931.2010.550505org/10.1080/02699931.2010.550505
53. McKnight, K. K. *et al.* Racial and regional differences in age at menopause in the United States: findings from the REasons for Geographic And Racial Differences in Stroke (REGARDS) study. *Am. J. Obstet. Gynecol.* **205**, 353.e1-353.e8 (2011).
54. Zeki, A. *et al.* Use of a pooled cohort to impute cardiovascular disease risk factors across the adult life course. *Int. J. Epidemiol.* **0**, (2018).