ESTIMATING THE IMPACTS OF INTERVENTIONS ON NON-AIDS RISK FACTORS IN OBSERVATIONAL HIV COHORTS

Tiffany Lyn Breger

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

Chapel Hill
2020

Approved by:

Stephen R. Cole

Jessie K. Edwards

Daniel Westreich

Brian W. Pence

Adaora A. Adimora

ABSTRACT

Tiffany Lyn Breger: Estimating the Impacts of Interventions on Non-AIDS Risk Factors in Observational HIV Cohorts
(Under the direction of Stephen R. Cole)

Non-AIDS risk factors contribute to persisting health disparities between people with HIV and the general population in the current era of effective combination antiretroviral therapy (ART). However, the optimal combination of interventions used in conjunction with ART to improve long-term outcomes remains unclear. In Aim 1, we used the parametric g-computation estimator to estimate the effects of combined interventions on non-AIDS risk factors on the risk of all-cause mortality among 1016 ART-naïve women enrolled in the Women's Interagency HIV Study (WIHS) between 1998 and 2017. Modeled interventions on alcohol and smoking combined with prompt initiation of modern ART decreased the 8-year risk of mortality compared to intervening solely on ART. Strategies that eliminated these non-AIDS risk factors achieved greater improvements in survival than strategies that reduced the prevalence of alcohol and smoking based on the expected efficacy of existing, real-world interventions.

In Aim 2, we developed and validated two-stage g-computation estimators that leverage partially observed information in the full study sample with complete exposure information available in a subset. Using a hypothetical cohort simulated to represent women with HIV enrolled in waves 2-4 of the WIHS, we illustrated a two-stage extrapolation g-computation estimator of the population average treatment effect and two-stage inverse probability weighted and exposure imputation g-computation estimators for the population average intervention effect. In 10,000 Monte Carlo simulations, two-stage approaches approximated the true values of the parameters of interest, considerably reduced bias and root mean squared error, and improved 95% confidence limit coverage compared to the naïve g-computation estimator fit to complete cases.

While modern ART has transformed the prognosis of HIV to a manageable chronic condition, non-AIDS risk factors remain significant contributors to early mortality. Our results suggest that interventions targeting alcohol and smoking may further reduce the risk of mortality. Achieving optimal health outcomes, however, will require more efficacious interventions as well as evaluation of interventions on other non-AIDS-related exposures that may not be completely measured in existing cohorts. While missing data threaten validity and precision, proposed two-stage g-computation estimators can be used to make progress in the face of these challenges.

To my mentors,
both professional and personal,
who helped me construct a wider lens
with which to view the world.

ACKNOWLEDGEMENTS

Writing this acknowledgement section of my dissertation is both the easiest and most difficult task of this whole process. While I was fortunate enough to receive an incredible amount of support from numerous individuals, it is not possible to fully convey the impact they have had on the completion of this body of work and my development as a scholar.

First, I would like to thank my committee. Steve Cole has been everything one could hope for in a mentor. He has challenged me, encouraged me, and advocated for me, and pushed me to think more deeply, question more often, and seek more opportunities to learn. One of the greatest gifts he has given me is teaching me how to learn. Jess Edwards has been nothing short of a rock star. My first encounter with Jess was when I attended her dissertation defense when I was visiting UNC as a prospective student. That became my vision of what doctoral level work looked like, and with no other reference or knowledge of what an extraordinary epidemiologist Jess was, I had no idea she had set the bar so high. In the four years I have been lucky enough to work with her, Jess has taught me to pursue creativity, think outside of the box, and never set limits on what is possible for myself. She has provided me with countless opportunities to grow and has exemplified the qualities of a mentor, scientist, and teacher that I hope to one day possess. Brian Pence has been with me from the very start. As my first advisor, he set the tone for my time at UNC and was the first person to recognize and believe in my capabilities when I did not see them. He set me along on my path as a methodologist and academic, a detour from the career in applied epidemiology that I had originally anticipated for myself. Daniel Westreich has been a constant source of inspiration for me and discussions with him always result in at least 10 new ideas. His enthusiasm in both my work and the future trajectory of the field of epidemiology always leaves me excited and ready for more. Ada Adimora has been an invaluable source of knowledge of the WIHS and

the research questions most pressing for people with HIV. She has helped me become a better communicator to clinical and health policy audiences and has kept me focused on the work that matters.

I am also appreciative of the funding I received throughout my doctoral training. The research for my dissertation papers was supported by the National Institutes of Health through the UNC WIHS (U01-AI-103390 PI: Adaora Adimora) and the Office of the Director and the Eunice Kennedy Shriver National Institute of Child Health & Human Development (DP2-HD-084070 PI: Daniel Westreich). Throughout my time at UNC, I was also fortunate to receive support from a UNC Center for AIDS Research supplement (PI: Jessie Edwards), the National Institutes of Health through the National Institute of Mental Health (R01-MH1000970 PI: Brian Pence; R01-MH087118-02 PI: Audrey Pettifor), and the African Studies Center as part of the Foreign Language and Area Studies Fellowship. I am incredibly fortunate to have always had a funding source and principal investigators that provided protected time to learn and advance my research interests.

My professors at UNC provided me with a solid foundation in this field. I especially thank those who were with me for the methods sequence – Steve Marshall, Daniel Westreich, Charlie Poole, Christy Avery, Brian Pence, Steve Cole, and Alan Brookhart. I must also thank Charlie Poole, with his nontraditional testing practices and personal anecdotes, for reshaping my mindset about learning and perceptions about making mistakes. I am not sure that I would have otherwise had the confidence or desire to remain in academia; nor would I have been a very good scientist. Throughout my year of grant writing, Steve Meshnick taught me the art of debate and writing a sales pitch that could not be refused.

My methods sequence TAs were phenomenal teachers and have become lifelong friends. Thank you Melissa Arvay, Anna Bauer, Liz Cromwell, Joann Gruber, Xiaojuan Li, and Laura McGuinn. Even semesters later, they have readily answered questions and provided expert advice long after their obligated commitment to me expired.

Somehow, I found myself in a number of exceptional study and research groups. My cohort was an amazing source of support and humor. I especially thank my fellow MSPHers – Devika Chawla, Nate DeBono, Brett Doherty, Lauren Levitz, Grace Mulholland, and Henry Zhang. The memories we made

together are some of my most cherished – not all of which could make it onto the Quote Wall. Thank you for the laughs, the tears, the celebrations, the late night delirium, the very bad epi jokes, the reminders to always check for missing semicolons, and the vague (or is it ambiguous?) "public health endeavors."

I am also grateful to the leaders of the MSPH seminar for their contribution in making us such a tight knit group – Nancy Colvin, Julie Daniels, Alan Kinlaw, and Steve Wing. It was in this setting that my first research ideas developed and were thoughtfully vetted. The conversations and philosophical discussions that were initiated in this venue during my first year are ones that I have come back to often. I thank Nancy for being a constant source of support and biggest advocate of the MSPHers and Alan for helping me navigate the epi program both during that first year and long afterward. I thank Julie and Steve for their mentorship. I also thank Steve for doing us all a great service and alleviating my perfectionist streak by not believing in "H's" and taking that off the table from day 1. But even more, I thank Steve for who he was as a person and scientist, for always asking tough questions, and for never providing the answers. As the years have gone by, I have missed Steve more and more, feeling his absence more significantly as I have progressed further through my training and have longed to return to some of those earlier conversations with him. I am fortunate for the time I had with him and for the continued lessons his words bring to me as time goes by. And maybe I have finally begun to understand why "confounding does not exist."

The Pence advisee group set an unbeatable standard for camaraderie – Kat Tumlinson, Julie O'Donnell, Nadya Belenky, Nalyn Siripong, Chris Gray, Sabrina Zadrozny, Aliza Liebman, Sara Levintow, Bryna Harrington, Jane Chen, Melissa Stockton, and Bethany DiPrete. Advisee group was always the highlight of my week and the place that I knew would always be filled with laughter. My first year at UNC, I felt like I had five additional advisors, benefiting from being the "youngest" of the group and getting to see my trajectory in all of you. I don't think there is any other advisee group out there that has more fun.

The Cole research group welcomed me in during my third year. Jen Lee, Sarah Willis, Terra Fatukasi, Jackie Rudolph, Haidong Lu, Dwight Yin, Catherine Li, and Rachael Ross – thank you for your

advice, questions, and contributions throughout my dissertation proposal and completion. My research is better because of you. I lost a wonderful office mate and IWHOD traveling companion in Jackie and hope our future career paths cross again.

Alex Keil and Alex Breskin also played key roles in my learning outside of the classroom and formal research group setting. Alex Keil taught me the g-formula and met with me numerous times over the years to answer all methods questions as well as to discuss general dissertation process and career advice. Alex Breskin was an amazing book club buddy and fellow methods enthusiast. He was the perfect person with whom to (attempt to) read Tsiatis and test out cool simulations.

Nancy Colvin, Valerie Hudock, and Jennifer Moore were the champions of the epidemiology department. Without meeting Nancy at APHA during my process of applying to grad schools, I very likely would not have ended up at UNC. Nancy is the reason I am here, and Nancy, Valerie, and Jennifer are the reasons I have stayed, managed to check all the boxes, and have done so without losing my sanity or well-being.

I am thankful to be surrounded and supported by such strong, intelligent, and powerful women working with me to revolutionize what it means to be women in science, academia, methods, and causal inference. Thank you to Chris Gray and Michele Jonsson Funk for dropping everything to help me plan the first Women in Epi town hall discussion in March 2017 and for encouraging me in making this an annual event and taking the lead on organizing various aspects. Thank you to Sabrina Zadrozny and Jess Edwards for being involved from the start and for continuing these conversations with me on a regular basis. And thank you to the women who have showed up, shared their experiences, and have given me hope and momentum for change.

"The Resistance" is what made Chapel Hill/Carrboro my home – Kat Tumlinson, Nadya Belenky, Chris Gray, and Sabrina Zadrozny. I entered UNC looking for an education and ended up with a family. They have been there for all of the highs and lows of graduate school and life – ready to share my frustrations, excitement, and hopes for the future and keep me going. I thank Kat for her love and kindness, for knowing what I needed and gently pushing me to get it, and for the thousands of novel-

length texts and emails checking in and helping me to sort things out even during the year that she was away. I thank Nadya for deciding to be my informal advisor, making sure I knew how to navigate my way through the program, and for all of the laughs. I thank Chris for being "the Chris Gray" – for showing me what it was to be a respected leader inside and outside of the department, making me believe that I had that potential myself, and doing more for me than I can recount. I thank Sabrina for all the early days spent working and laughing at Weaver, pushing roadblocks out of the way (both literally and figurately), and for always being intentional in making me feel like I belonged.

And through The Resistance, I also gained a larger bonus family. Mehul Patel readily accepted me as a constant in Kat's life and became one of my closest and most trusted friends himself. Somehow, he always knows the right thing to say. The Grays have given me some of the best holiday memories and have been some of my biggest cheerleaders. The Patels have incorporated me as a member of their family and have shown me so much love and support. I also must thank the world's greatest nephews and my most favorite redheads – Oliver, Arthur, and Peanut. They are my whole heart and the most precious beings in my life.

Finally, I am deeply indebted to the participants of the Women's Interagency HIV Study without whom this research would not be possible. It is they who made the ultimate commitment to science and the greatest contribution to this work. They dedicated their time, energy, and personal histories to benefit the future of others. At the beginning of my first semester at UNC, Steve Marshall paraphrased a quote that's been attributed to a number of early health researchers – "Statistics are people with the tears wiped away." This is something that can be easy to forget or fully appreciate while working with data year after year, especially when caught in the weeds of methodological nuances. But there is so much more there that cannot be seen in the necessary reduction of each individual to a data point for analysis. These are real human beings with their own individual experiences, memories, families, and hopes for their lives and the lives of others. It is the responsibility of the researcher to prove themselves worthy of these individuals' trust and to not let their sacrifices have been made in vain. I thank the women for being the real heroes of this work.

TABLE OF CONTENTS

LIST OF TABLES xiv

LIST OF FIGURES xv

LIST OF ABBREVIATIONS xvi

CHAPTER 1: INTRODUCTION 1

A. Moving Beyond Antiretroviral Therapy: Intervention Portfolios Among People with HIV 1

B. Development and Assessment of Novel Statistical Methods: A Brief History 6

CHAPTER 2: STATEMENT OF SPECIFIC AIMS 9

A. Overview 9

B. Aim 1 9

C. Aim 2 10

CHAPTER 3: METHODS 12

A. Overview 12

A1. Parametric G-computation Estimator of the Parametric Generalized-Formula 12

A2. Identification Conditions 13

B. Aim 1: Interventions on Non-AIDS Risk Factors among Women with HIV 14

B1. Data Source and Study Population 15

B2. Study Inclusion and Exclusion Criteria 16

B3. Exposures and Outcome 17

B4. Interventions 18

B5. Covariates of Interest 20

B6. Treatment Decision Design 22

B7. Notation and Parameters of Interest 23

B8. Implementation of Parametric G-computation 26

C. Aim 2: Development and Implementation of Two-Stage G-computation Estimators 28

C1. Motivating Example 29

C2. Study Data 30

C3. Estimators 32

C4. Statistical Analyses 36

CHAPTER 4: ALL-CAUSE MORTALITY UNDER MODELED INTERVENTIONS ON ANTIRETROVIRAL THERAPY, ALCOHOL, AND SMOKING AMONG HIV-POSITIVE WOMEN IN THE UNITED STATES, 1998 – 2017 39

A. Overview 39

B. Introduction 40

C. Methods 40

C1. Study Sample 40

C2. Exposures and Outcome 41

C3. Intervention Portfolios 42

C4. Analyses 43

D. Results 44

E. Discussion 45

F. Conclusion 49

CHAPTER 5: TWO-STAGE G-COMPUTATION: EVALUATING TREATMENT AND INTERVENTION IMPACTS IN OBSERVATIONAL COHORTS WHEN EXPOSURE INFORMATION IS PARTLY MISSING 55

A. Overview 55

B. Introduction 55

C. Methods 57

C1. Motivating Example Data 57

C2. G-computation 58

C3. Two-stage G-computation Estimator of Average Treatment Effect 60
C4. Two-stage G-computation Estimators of Average Intervention Effect 60
C5. Analyses of Example Cohort 62
C6. Simulation Experiments 63
D. Results 63
D1. Motivating Example Cohort 63
D2. Simulation Experiments 64
E. Discussion 65
F. Conclusion 68
CHAPTER 6: DISCUSSION 80
A. Overview of Key Findings 80
A1. Aim 1 80
A2. Aim 2 82
B. Limitations 83
C. Strengths 88
D. Summary and Public Health Significance 92
APPENDIX A: SUPPLEMENTAL TEXT 95
APPENDIX B: SUPPLEMENTAL TEXT 106
REFERENCES 111

LIST OF TABLES

Table 4.1. Characteristics of 1,016 HIV-positive combination antiretroviral therapy (ART) naïve women in the Women's Interagency HIV Study (WIHS) at analysis baseline. 50

Table 4.2. Risk of all-cause mortality under hypothetical intervention portfolios to eliminate non-AIDS risk factors 51

Table 4.3. Risk of all-cause mortality under hypothetical intervention portfolios to reduce the prevalence of non-AIDS risk factors 51

Table 5.1. Observed and complete data for a simulated cohort of 1,623 HIV-positive women. 73

Table 5.2. Estimates of the twelve-month risks (95% confidence intervals)[a] of emergency room visit or death 73

Table 5.3a. Performance of g-computation approaches addressing missing data to estimate average treatment effect of shortening versus lengthening opioid prescription duration 74

Table 5.3b. Performance of g-computation approaches addressing missing data to estimate average intervention effect of shortening opioid prescription duration relative to the status quo 75

Table 5.4a. Performance of g-computation estimators of the average treatment effect under varied sample conditions with 30% complete opioid information. 76

Table 5.4b. Performance of g-computation estimators of the average intervention effect under varied sample conditions with 30% complete opioid information. 78

LIST OF FIGURES

Figure 3.1. Women's Interagency HIV Study and data coordinating center sites. 15

Figure 3.2. Causal Diagram of ART, Non-AIDS Risk Factors, and Mortality. 21

Figure 4.1. Risk of all-cause mortality under four antiretroviral therapy initiation strategies in the Women's Interagency HIV Study, 1998 – 2017. 52

Figure 4.2. Risk of all-cause mortality under hypothetical interventions to eliminate (Panel A) or reduce (Panel B) the prevalence of non-AIDS risk factors in the Women's Interagency HIV Study, 1998 – 2017. 53

Figure 4.3. Observed versus g-computation simulated cumulative incidence of all-cause mortality under the natural course of antiretroviral therapy initiation 54

Figure 5.1. Causal diagram used to generate a hypothetical cohort of 1,623 HIV-positive women 69

Figure 5.2. Risk difference estimated by g-computation approach contrasting provision of short duration versus long duration opioid prescriptions 70

Figure 5.3. Bias (Panel A) and standard error (Panel B) for the average treatment effect of shortening versus lengthening opioid prescription duration 71

Figure 5.4. Bias (Panel A) and standard error (Panel B) for the average intervention effect of shortening opioid prescription duration versus the status quo prescribing practice 72

Appendix Figure 1. Risk difference for the average treatment effect estimated by g-computation approaches including and excluding covariate Z. 109

Appendix Figure 2. Bias (Panel A) and standard error (Panel B) for the effect of shortening opioid prescription duration among 7,500 HIV-positive women 110

LIST OF ABBREVIATIONS

AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
CI	Confidence Interval
IPTW	Inverse Probability of Treatment Weighting
IPW	Inverse Probability Weighting
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
IDU	Injection Drug Use
RD	Risk difference
RMSE	Root mean squared error
RR	Risk ratio
US	United States
WIHS	Women's Interagency HIV Study

CHAPTER 1: INTRODUCTION

A. Moving Beyond Antiretroviral Therapy: Intervention Portfolios Among People with HIV

Health disparities persist between people with human immunodeficiency virus (HIV) and the general population. Though acquired immunodeficiency syndrome (AIDS)-related mortality has dramatically declined in the United States (US) since the availability of highly effective combination antiretroviral therapy (ART), other causes of morbidity and mortality disproportionately affect those with HIV.[1–4] Compared to the general US population, those with HIV experience a higher prevalence of comorbidities including hypertension,[5] cardiovascular disease,[6–9] diabetes,[5,10] liver disease,[11,12] and pulmonary disease.[13–15] Up to 65% of people with HIV experience the co-occurrence of 2 or more chronic conditions (i.e., multimorbidity).[16] In addition to this high burden of comorbidities, people with HIV experience accelerated aging with an onset of multimorbidity occurring a decade earlier than those without HIV.[17–19] Consequently, the widespread adoption of ART has not closed the gap in survival for people with HIV who continue to die from non-AIDS causes 8 to 9 years earlier than those without HIV.[20–22]

Further improvements in the health status and survival of people with HIV will require greater attention to non-AIDS risk factors. Cigarette smoking, substance use, and chronic inflammation are highly prevalent among HIV-positive populations and may contribute to a larger proportion of excess morbidity and mortality in the modern ART era. In a nationally representative US survey, over 40% of adults with HIV reported current smoking, compared to 20% of the general population.[23] The proportion of adults with HIV reporting any alcohol consumption in the last month has ranged from 40-53% in US based cohorts, with an estimated prevalence of heavy drinking twice that of the general population.[24–27] Additionally, chronic inflammation persists even after successful viral suppression and might increase the

risk of cardiovascular disease as people live longer with HIV.[28–30]

Cigarette smoking and alcohol are important risk factors for many chronic conditions in the general population. Both smoking and heavy alcohol use have been associated with elevated risks of hypertension,[31] cardiovascular disease,[31] and liver disease.[32,33] Furthermore, smoking and alcohol increase inflammation, which is a known marker of early mortality.[34–36] In addition to cigarette smoking being an established risk factor for pulmonary diseases,[37,38] it has also been identified as a risk factor for Type 2 diabetes.[39,40]

While tobacco, alcohol, and inflammation are recognized as significant risk factors in the general population, their adverse effects may be more pronounced among those with HIV. Several studies have reported a higher risk of respiratory tract infections, chronic obstructive pulmonary disease, and lung cancer among smokers with HIV, suggesting a synergistic relationship between cigarette smoking and HIV infection.[15,41–46] In the Veterans Aging Cohort Study, HIV status appeared to interact with alcohol consumption in increasing the risk of adverse health outcomes among men; those with HIV experienced an increased risk of physiologic injury and mortality at lower thresholds of alcohol consumption than their HIV-negative counterparts.[47] With people now less likely to die of AIDS and living longer with HIV in the modern ART era, these non-AIDS risk factors are becoming important determinants of excess morbidity and mortality.

Though there have been calls for a more comprehensive approach to HIV care management, several limitations to the evidence base have impeded public health action. First, existing observational studies have primarily focused on quantifying the magnitude of the relationship between specific non-AIDS risk factors and mortality. Considerable effects have been reported in select populations, including a two-fold increase in mortality with smoking,[42] a 40% higher hazard of mortality associated with heavy alcohol consumption,[27] and a 2.6-2.7-fold higher odds of mortality associated with fibrinogen and C-reactive protein inflammatory markers.[48] However, it is not possible to infer from these studies the potential reduction in mortality that would be achieved by interventions that combine prompt initiation of modern ART with the elimination or reduction of the prevalence of these risk factors relative to the status

quo.[46,49–51] The estimated effect of a risk factor on an outcome cannot be interpreted as the effect of intervening on that risk factor given potential unmeasured, indirect effects of the intervention on the outcome and imperfect intervention efficacy.[52]

While randomized controlled interventions targeting non-AIDS risk factors have demonstrated some promising results they have largely been conducted among those without HIV. Immunocompromised health status is a common exclusion criteria in trials which has led to an existing evidence base that is more applicable to relatively healthy, non-pregnant adult populations. In a meta-analysis of smoking cessation randomized interventions, nicotine replacement therapy treatments, bupropion, varenicline, and cytisine demonstrated efficacy in quitting rates.[53] However, these interventions may have differing effects among those with HIV, or in some cases, be inadvisable. For example, in several studies varenicline has been associated with increased psychiatric symptoms among those with depression[54] – a comorbidity that is extremely common among those with HIV.[55] Furthermore, pharmacological based interventions may not be appropriate for women of reproductive age, particularly if pregnancy is likely. The U.S. Preventive Services Task Force has declared that there is insufficient evidence to evaluate the benefits versus harms of pharmaceutical smoking cessation aids during pregnancy.[56]

Most alcohol consumption interventions have also been conducted among non-generalizable study populations. Some cognitive-behavioral therapy and pharmacological-based interventions have demonstrated efficacy in short-term reductions in alcohol consumption as well as lower rates of relapse.[57–59] However, studies have been conducted primarily among those with alcohol dependence or alcohol consumption levels far beyond the national recommended safe limits.

Studies of similar interventions among people with HIV have either been absent or enrolled small, highly selective (often majority male) samples of HIV-positive individuals and evaluated only short-term effects. In behavioral smoking cessation trials,[60] interventions have been highly variable and endpoints have typically been smoking cessation at 6 or 12 months, making it difficult to draw conclusions about the long-term real-world effectiveness of various strategies. Trials evaluating the

efficacy of interventions targeting alcohol use have been conducted among high risk groups and yielded mixed results;[61] no studies have evaluated the impact of lowering alcohol consumption to national recommended limits among the majority of those who consume alcohol but are not classified as heavy drinkers. Similar to the challenge of randomized controlled trials outside of HIV settings, women are often underrepresented in trials of interventions on non-AIDS-related risk factors.

Further evidence of the real-world effectiveness of existing interventions among those with HIV is necessary to guide clinical decision making. However, even in observational HIV cohorts in which collected exposure and risk factor information may enable a more accurate depiction of the experience of those with HIV in the US and improve target validity,[62] women are underrepresented. Many existing studies have been restricted to men or are conducted in cohorts that include an overwhelming majority of men. The Veteran's Aging Cohort Study is composed of 97% men which made it necessary to restrict to men while studying the interaction between HIV and alcohol consumption on physiologic injury and mortality.[47] No analogous study has been conducted among women with HIV. Other large cohorts of HIV-positive individuals in the US, such as the CFAR Network of Integrated Clinical Systems, are composed of over 80% men.[63] Thus, results from studies conducted in these cohorts are not necessarily generalizable to women with HIV who experience a different distribution of risk factors than men with HIV[5] and may also experience differing intervention efficacy. Yet women with HIV experience far greater multimorbidity[5] and have seen less improvement in life expectancy with ART than men with HIV[20] which makes it imperative to study multiple interventions on non-AIDS risk factors in this group using information collected in HIV cohorts that have enrolled large samples of women.

Few studies to date, however, have looked at effects of sets of interventions or compared interventions targeting different risk factors in a given population of HIV-positive individuals either generally or among HIV-positive women specifically. In both randomized trials and observational cohort settings, researchers have often focused on estimating the effect of a particular intervention in isolation, or less frequently, the effect of a pair of interventions.[64,65] However, given the clustering of several risk factors and comorbidities among those with HIV (especially women), information about more

comprehensive intervention strategies is important. It is possible that improvements in survival may require more than one intervention. For example, those who quit smoking but still engage in regular, heavy drinking may not experience a substantial delay in mortality that would otherwise be observed by intervening on both risk factors together. Synergistic relationships between interventions might also exist, with early intervention on more than one risk factor yielding a greater improvement in clinical outcomes than intervening on each risk factor separately. Comparisons of intervention combinations are critical to clinicians and policymakers in prioritizing strategies for comprehensive HIV care management.

Traditional study approaches cannot provide clinicians and policymakers with estimates of the potential improvements in outcomes among those with HIV that can be achieved by various combined interventions. Ideally, a factorial randomized trial would be conducted to evaluate sets of interventions on multiple risk factors. In light of constrained resources, the large sample size and length of follow-up period required preclude these studies from being conducted. While observational studies offer a plethora of longitudinal data on an often less selective study population, confounding is a major concern due to nonrandomized treatment. Additionally, standard approaches using multivariable regression models to address confounding fail to provide valid estimates in the presence of time-varying confounders[66] and often do not give covariate marginal effects.

Recent advances in statistical software and causal inference methods offer a unique opportunity to leverage observational data to quantify the impact of potential intervention portfolios (i.e., combinations of interventions). Marginal structural models, most commonly estimated with inverse probability of treatment weights, can appropriately control time-varying confounding and provide estimates of causal effects when sufficient conditions are met.[66] However, inverse probability of treatment weighting is not always conducive to providing estimates of intervention contrasts,[49,51] especially when multiple, combined interventions are being assessed. The parametric g-computation estimator is a particularly flexible estimator of the parameters of a marginal structural model which can be used to estimate population-level effects of various changes to exposure distributions.[49,51,67,68] Because the estimator is fully parametric, it can achieve improved efficiency compared to inverse probability

weighting in some settings.[64,68] Recent illustrations of its application using available statistical software packages have now made it a viable analytic tool.[46,65,68–75]

B. Development and Assessment of Novel Statistical Methods: A Brief History

Causal inference methods are increasingly applied to observational data to answer pressing research questions and guide clinical decision making. For both HIV-positive populations and other vulnerable groups, there is an urgency to identify and rigorously evaluate interventions that can improve patient health outcomes. Randomized controlled trials are costly, inefficient, and infeasible in many settings. Consequently, causal inference methods play an important role in providing answers to these questions. Examples have demonstrated that under a set of "identification" assumptions, these methods can be applied to existing observational data to yield results comparable to those that would be obtained with a randomized controlled trial.[66,68,75,76]

The parametric g-computation estimator is one such causal inference method that is likely to become a more popular approach to assessing potential impacts of interventions. Until recently, a major barrier to the adoption of the parametric g-computation estimator was the absence of examples of its implementation using available statistical software. However, a PubMed search reveals that since Taubman et al. provided an illustration in 2009,[69] there have been 75 published studies using this estimator in analyses. Given the unique flexibility of the parametric g-computation estimator and additional guidance published in 2012[68] and 2014,[71] it will likely become a more common analytic choice.

Approaches to implement the parametric g-computation estimator in analyses assume complete exposure information; yet, missing data are extremely common in epidemiologic research.[77] In previous studies using HIV observational data, this has not presented a substantial problem. This is due to the fact that in many instances to date, the research question has focused on ART based interventions.[68,73,78] Because many existing observational HIV cohorts were established to study the natural and treated progression of HIV, ART information tends to be very well measured with few missing data.

However, with persisting health disparities between ART-treated people with HIV and the general population, attention is shifting to evaluating impacts of interventions on non-AIDS risk factors.

Further improvements in the health outcomes and survival of people with HIV will require more comprehensive interventions that additionally target non-AIDS risk factors; yet, sparsely collected data on non-ART exposures in existing observational data sources leaves investigators without validated approaches to quantify such intervention effects.

In both existing HIV cohorts as well as other observational studies, exposure information may be missing by design or missing for unknown reasons. Missing data may occur by design when data collection is expensive or was only performed on a subset of the cohort. For example, the Women's Interagency HIV Study has additional detailed biomarker and behavioral data collected on a sample of cohort members as part of previous and ongoing substudies. (See: http://statepiaps.jhsph.edu/wihs/invest-info/dossier.pdf) Additionally, linkages to claims databases or electronic medical records can provide more detailed exposure information but may only be possible for a subset of the cohort due to cost or other barriers such as only being able to link records for people who are in regular care. To avoid loss of efficiency and potential threats to the validity of results,[77,79–81] investigators need guidance regarding approaches that may be used to address missing data in the parametric g-computation framework.

Neyman introduced two-stage study sampling designs as an approach to improve accuracy of estimators when information on a particular variable may be difficult or costly to measure among the total population of interest.[82] "Two-stage" study designs have since been illustrated using standard regression models to recover valuable information from partially observed cases, particularly in the context of case-control and case-cohort data.[80,83–87] While such study designs have improved the accuracy of regression coefficients estimated under various missing data structures, these parameters do not approximate the population-level intervention effects needed to inform policy.[49–51,88] Importantly, existing two-stage methods have not been adapted to compare counterfactual outcome distributions using the parametric g-computation estimator when exposure is partially missing. As a result, parametric g-computation estimator analyses have been limited to 1) address questions in which exposure information is always complete, limiting the types of questions that can be answered (particularly in regards to non-AIDS risk factors); 2) include only individuals with complete exposure information, reducing external validity of

findings; or 3) use ad-hoc methods to account for missing data which have not been systematically validated for g-computation.

New extensions to rigorous counterfactual-based approaches will be imperative to leverage underused observational data to evaluate and compare interventions on non-AIDS risk factors. Furthermore, such approaches are necessary to guide efficient, cost-effective design of future studies that have the potential to improve healthcare delivery strategies and ultimately reduce the burden of morbidity and early mortality among those with HIV.

CHAPTER 2: STATEMENT OF SPECIFIC AIMS

A. Overview

The overall objectives of this work were to 1) leverage observational data and novel quantitative methods to estimate the impact of combined interventions to delay mortality among women with HIV and 2) refine these methods to accommodate missing exposure information that is likely to hamper crucial future studies of interventions on non-AIDS risk factors. To meet these objectives, we carried out the following specific aims using longitudinal data from HIV-positive women participating in the Women's Interagency HIV Study (WIHS)[89–91]– the largest ongoing interval cohort of HIV-positive women in the United States.

B. Aim 1

We aimed to estimate the effects of alcohol consumption and smoking cessation interventions combined with prompt initiation of ART in the modern treatment era on the 8-year cumulative incidence of all-cause mortality in the WIHS between 1998 and 2017 using the parametric g-computation algorithm. Our goal was to use this approach to estimate the risks of all-cause mortality under two sets of intervention portfolios (i.e., combined interventions) where Set A interventions were specified as the elimination of these non-AIDS risk factors and Set B interventions were specified as the reduction in the prevalence of these non-AIDS risk factors based on the expected efficacy of existing, real-world interventions. We further aimed to estimate the impact of each intervention portfolio compared to an intervention only on prompt initiation of ART in the modern era using risk differences and ratios.

Hypothesis: We hypothesized that interventions on non-AIDS risk factors in addition to prompt initiation of ART would reduce the risk of all-cause mortality and that interventions that eliminated,

rather than reduced the prevalence of, non-AIDS risk factors would achieve the greatest improvements in survival.

Rationale: There is a high prevalence of smoking and alcohol consumption in the WIHS[91–94] and other HIV-positive populations,[23,24,95] both of which are known risk factors for mortality and have potentially more pronounced adverse effects among those with HIV.[15,41–45,47] With ART-treated HIV-positive individuals now less likely to die from AIDS, these non-AIDS risk factors may contribute to a greater proportion of excess morbidity and mortality. Consequently, eliminating or reducing the prevalence of smoking and alcohol consumption may improve survival in the modern ART era. Because eliminating risk factors assumes interventions with 100% efficacy and no side effects[52] and existing interventions have much lower efficacy, the smaller reduction in the prevalence of smoking and alcohol consumption with real-world interventions will likely have attenuated effects on survival.

C. Aim 2

We aimed to develop, validate, and illustrate approaches to constructing two-stage g-computation estimators of the total population-level treatment and intervention effects in the presence of missing exposure information. Using a simulation study design and a hypothetical cohort simulated to represent HIV-positive women enrolled in the WIHS, our goal was to validate the following two-stage estimators. For the estimator of the average treatment effect, our goal was to construct and validate a two-stage extrapolation approach in which conditional probabilities are estimated from parametric models fit to a subset of study participants with complete exposure data; in the second stage, conditional probabilities are directly extrapolated to the full cohort. For the estimators of the average intervention effect, our goal was to construct and validate two-stage inverse probability weighted g-computation and exposure imputation g-computation approaches. We further aimed to compare the performance of each of the proposed estimators in terms of bias, standard error, 95% confidence limit coverage, and root mean squared error.

Hypothesis: We hypothesized that when exposure information is not missing completely at random, the naïve g-computation algorithm fit to complete cases will provide biased estimates of absolute and relative risks in the population of interest; however, the two-stage extrapolation g-computation

estimator will provide unbiased estimates of the absolute and relative risks for an "always treated" and "never treated" population, provided all covariates predicting both missingness and the outcome are included in models. Similarly, the two-stage inverse probability weighted and exposure imputation g-computation estimators will provide unbiased estimates of risks for an "always treated" population compared to the population under the "natural treatment course" (i.e., no imposed intervention).

Rationale: When exposure information is missing not at random, absolute and relative risks estimated from those with complete information will not necessarily equal those that would have been observed among the full cohort of interest. Two-stage g-computation approaches that leverage available covariate and outcome information from all individuals, including those with missing exposure information, will improve precision and consistency, as has been illustrated in other analytic settings.[80,83,84,96] Because the inverse-probability weighted and exposure imputation g-computation approaches include a model for the exposure, the outcome distribution under the natural treatment course can be recovered.

CHAPTER 3: METHODS

A. Overview

This chapter presents the study design, data source, assessment of measures, and statistical methods as they relate to Aims 1 and 2. The chapter is organized as follows. Section A presents a brief overview of the parametric g-computation algorithm – the approach that is used in Aim 1 analyses to estimate the effects of interventions on non-AIDS risk factors and extended in Aim 2 to accommodate missing exposure information when estimating treatment and intervention effects. In Section B, we outline the study sample, variable measurement, and analyses for Aim 1. In Section C, we outline the study design, developed estimators, and analyses for Aim 2.

A1. Parametric G-computation Estimator of the Parametric Generalized-Formula

In this document, we make the following distinction between the terminology "g-formula" and the "parametric g-computation estimator." We use "g-formula" to refer to an equation that is used to express the observed data distribution as the distribution of the data that would have been observed under an alternative treatment plan. We use "parametric g-computation estimator" to refer to a parametric estimator of the g-formula equation. We also use "algorithm" interchangeably with "estimator."

The parametric g-computation formula estimator[67] is a recently illustrated approach[68,69,71,78,97–99] to estimating the parameters of a marginal structural model. Analogous to an inverse probability of treatment weighted (IPTW) estimator of the parameters of a marginal structural model,[66,76,100] the g-computation estimator is a generalization of standardization. In longitudinal observational study settings subject to time-varying confounding, both IPTW and g-computation estimators can provide valid estimates of the total treatment effect that would otherwise be biased with the use of traditional regression estimators.[67,101]

Three particular advantages of the parametric g-computation estimator are its flexibility in estimating population-level effects of realistic changes to exposure distributions,[46,49,51,65,74,102,103] ability to estimate effects of interventions that depend on the natural value of exposure,[104,105] and efficiency in quantifying impacts of multiple interventions.[69,105] Whereas the IPTW estimator has been widely used to obtain contrasts of an "always treated" versus "never treated" population, this is often not representative of the scenarios that policymakers are considering.[49,51] For example, in deciding whether more resources should be allocated to smoking cessation interventions, the alternative to targeting more smokers for treatment is not "take away treatment from all smokers"; rather, it is more likely that the alternative is "maintain the status quo." The parametric g-computation estimator can easily be specified to provide such contrasts. The extended parametric g-computation estimator can also be used to specify interventions that involve changes to exposure that depend on the natural value of exposure that would be observed if intervention were discontinued immediately before measurement.[104,105] An example of an intervention that depends on the natural value of exposure is "intervene to eliminate depression with probability 66% if depression occurs without intervening immediately before this timepoint."[64] Finally, while the IPTW estimator can become inefficient when estimating multiple joint interventions, the g-computation estimator relies on more parametric models, and consequently, may maintain a higher level of efficiency.[64,106]

A2. Identification Conditions

Under a sufficient set of conditions, the g-computation formula re-expresses the distribution of the data under the observed treatment (i.e., the factual) as the distribution of the data that would have been observed under a potentially alternative treatment (i.e., the counterfactual). Mathematically, this can be written in a simplified, time-fixed setting as shown below where the left-hand side of the equation represents the observed expectation, the right-hand side represents the counterfactual expectation, random variables are denoted with uppercase letters, realizations of those variables are denoted with lowercase letters, the outcome is denoted Y, the treatment is denoted A, and covariates are denoted W.

$$\sum_{w} E(Y = y|A = a, W = w)P(W = w) = E\{Y(a) = y\}$$

The identification conditions are as follows:

1) No measurement error of treatment, covariates, or outcome.[107] More specifically, we assume that the treatment and outcome are measured without error while covariates might be mismeasured but that the observed covariate values are those that inform the treatment plan. For example, a laboratory measurement of CD4 cell count might be 220 for a patient whose true CD4 cell count is 235; however, the 220 value the clinician sees is the value that is used to determine whether or not to initiate ART.
2) Counterfactual consistency (also described as treatment version irrelevance and no interference).[108,109] If the observed treatment of patient *i* is equal to the treatment of interest (i.e., $A_i = a$), then the observed outcome of patient *i* is equal to the outcome that would have been observed if treatment was set to the treatment of interest (i.e., $Y_i = Y_i(a)$).
3) Conditional exchangeability or no unmeasured confounding.[110] The probability of treatment conditional on a set of covariates is independent of the potential outcome (i.e., $P(A = a|W = w, Y(a)) = P(A = a|W = w)$).
4) Positivity, also described as the experimental treatment assignment assumption.[111] The conditional probability of receiving treatment is nonzero within all strata of covariates (i.e., $P(A = a|W = w) > 0$).
5) Correct specification of parametric models.

B. Aim 1: Interventions on Non-AIDS Risk Factors among Women with HIV

The overall goal of Aim 1 was to use the parametric g-computation estimator to quantify the impact of combined interventions on non-AIDS risk factors and prompt initiation of modern ART on the risk of all-cause mortality in the Women's Interagency HIV Study.

B1. Data Source and Study Population

The Women's Interagency HIV Study (WIHS) is an ongoing, multisite prospective cohort study of HIV-positive and HIV-negative women in the US.[89–91] Though the WIHS was initially established in 1993 to study the progression of HIV infection in women,[89] its research focus has since expanded to other areas including aging, behavioral and social health determinants, comorbidities, and epidemiologic methods.[91] As shown in Figure 3.1, the WIHS includes 10 consortia of contributing sites, 6 of which began in 1994 (Bronx/Manhattan, NY; Brooklyn, NY; Chicago, IL; Los Angeles/Southern CA/Hawaii; San Francisco/Bay Area, CA; Washington, DC) and 4 of which were added in 2013 (Atlanta, GA; Birmingham, AL/Jackson, MS; Chapel Hill, NC; Miami, FL). The Baltimore, MD site has served as the data coordinating center since 1997.

Figure 3.1. Women's Interagency HIV Study and data coordinating center sites.

Source: https://wihs.gumc.georgetown.edu/aboutwihs/partnersites

Women with HIV and women at risk of acquiring HIV were enrolled during 4 recruitment waves (Wave 1: 1994-1995; Wave 2: 2001-2002; Wave 3: 2011-2012; Wave 4: 2013-2015) and have since been followed at visits every 6 months. Women were recruited from a wide range of clinical, research, and community outreach sites which has improved representation of women with HIV, especially those who are not in regular HIV care and are harder to reach.[89] Because of its recruitment strategies, the WIHS has an age and racial/ethnic distribution that is representative of women with HIV in the US;[91] it is the largest

and longest ongoing prospective cohort study of women with HIV in the US. In other large clinical cohorts of people with HIV in the US, women (especially non-White women) are a small minority, limiting generalizability among this population. Thus, the WIHS is one of the leading cohorts for studies among women with HIV in the US.

WIHS participants attend study visits every 6 months, during which they contribute extensive demographic, medical history, medication use, health service utilization, and behavioral data through in person interviews. During visits, participants also undergo clinical examination including vital sign measurements, anthropometric measurements, and a gynecological exam. Blood, urine, and cervicovaginal swab samples are collected. Data from study visits are supplemented with medical record abstractions and linkages to cancer and death registries.

B2. Study Inclusion and Exclusion Criteria

Since 1994, the WIHS has enrolled 4982 women, 3703 of whom were HIV-positive at baseline or seroconverted over follow-up. Eligibility for Aim 1 of the proposed study was restricted to HIV-positive women with a study visit on, or after, 1 April 1998 and who had not previously initiated ART (N=1033). The analysis baseline for each woman was defined as the woman's first HIV-positive WIHS study visit on, or after, 1 April 1998. For women enrolled in Wave 1 (1994-1995), the analysis baseline date occurred after the date at which they entered the WIHS. For women enrolled in all other waves, the analysis baseline date was equal to the WIHS entry date if they were HIV-seropositive at entry; otherwise, it occurred after their WIHS entry date if they seroconverted over follow-up.

We defined the baseline visit as occurring on, or after, 1 April 1998 because our research questions were relevant to the calendar period in which highly effective, triple therapy combination ART was widely available. It was after the introduction of these regimens in 1997[112] that AIDS-related morbidity and mortality most significantly declined. Thus, we focused on the calendar period from 1998 onward since this is when women initiating ART would be most likely to initiate triple combination regimens as opposed to less effective 2-drug therapy. The beginning of 6-month intervals for WIHS study

visits occur on 1 April and 1 October each year – hence, the inclusion of visits occurring at/after 1 April 1998. All follow-up visits after each woman's analysis baseline date were included.

Women who had already initiated 3-drug combination ART prior to the analysis baseline date were not included. This was to accommodate a new-user design in regards to ART treatment.[113] In observational study settings, new-user designs are recommended to eliminate biases that arise when prevalent users are included in analyses. Namely, new-user designs prevent survivor bias in regards to prevalent users that survive early periods of pharmacotherapy and bias resulting from confounders that are affected (i.e., mediated) by the treatment.

Among 1033 HIV-positive combination ART-naïve women with a study visit on/after 1 April 1998, we excluded 2% of the sample for missing information on key baseline variables. We excluded women with missing baseline information on depression (n=17), alcohol consumption (n=9), and smoking status (n=8). In the remaining sample of 1016 HIV-positive women, we imputed 5 (0.5%) missing baseline CD4 counts and 18 (1.8%) missing baseline viral load counts by simulation from log normal distributions with mean and standard deviation equal to that of the observed WIHS sample at baseline. The final sample for analysis included N=1016 HIV-positive, combination ART-naïve, women.

B3. Exposures and Outcome

Information on ART, alcohol consumption, and smoking exposures have been collected on WIHS participants via in-depth interviews at all baseline and follow-up visits. Combination ART is determined using a definition based on Department of Health and Human Services/Kaiser Panel guidelines.[114] Women are considered to be on combination ART if they report use of 3+ antiretroviral drugs, one of which is a protease inhibitor (PI), a non-nucleoside reverse transcriptase inhibitor (NNRTI), an integrase inhibitor (II), or an entry inhibitor (EI). During interviews, women also reported the number of drinks consumed on average per week and smoking status since the last visit.

The WIHS has confirmed deaths through medical records, state death certificates, and the National Death Index Plus (NDI+) for all women who have been enrolled in the WIHS. State death certificate data are available from 1994 until state death registry queries were discontinued in 2015. NDI+

data are available from 2000 until present. In both state and national death registries, data include the date and location of death as well as the International Classification Diseases (ICD) codes for the underlying and contributing causes of death. Our outcome of interest was all-cause mortality and was defined as a documented death from any cause.

B4. Interventions

Our primary objective was to estimate the risk[115,116] of all cause-mortality in the WIHS under hypothetical intervention portfolios (i.e., intervention combinations) targeting prompt initiation of modern ART, alcohol, and smoking. In our base scenario (Portfolio 1), we considered an intervention solely on prompt initiation of modern ART to represent the "status quo" in regards to the current recommended healthcare guidelines among people with HIV. We defined modern ART as combination ART initiated on/after 1 October 2001. This date coincides with the period after which tenofovir had been approved and became a common drug in combination ART regimens.[114] "Prompt" initiation was defined as initiation of ART by 6 months of baseline (i.e., being on ART at the first follow-up visit). This definition was influenced by the structure of the WIHS as an interval cohort with exposure status documented at the time of each 6-month follow-up visit. Because ART initiation dates are documented as the first WIHS study visit date at which an individual reported being on ART, it was not possible to decipher the exact point at which the individual initiated ART in the period between the last ART-naïve study visit and the next follow-up.

In our other three scenarios, we considered portfolios combining prompt initiation of modern ART with: an intervention on alcohol (Portfolio 2), an intervention on smoking (Portfolio 3), and interventions on both alcohol and smoking (Portfolio 4). Specifically, we defined interventions on non-AIDS risk factors in two ways. Set A intervention portfolios were those that combined prompt initiation of modern ART with *near elimination* of alcohol consumption and/or smoking whereas Set B portfolios were those that combined prompt initiation of modern ART with *realistic reductions* in the prevalence of alcohol consumption and/or smoking based on the expected efficacy of real-world interventions.

Set A intervention portfolios were defined as follows. Portfolio 2A was an intervention to eliminate alcohol consumption of over 1 drink per week among women without an indication of hepatitis C virus (HCV) and eliminate all alcohol consumption among women with an indication of HCV. HCV status was determined based on bloodwork performed on all WIHS participants at baseline. If women were positive for HCV antibody, they received testing for HCV RNA to determine whether the infection was active. Here, we defined an indication of HCV to be a positive antibody test and a positive or missing HCV RNA test. We chose a limit of 1 drink per week instead of complete abstinence among those without HCV due to potential protective effects of alcohol at low levels[117] but did not define the intervention to increase drinking among those who were already abstainers. Because alcohol is particularly harmful for those with hepatitis C infection,[118] the intervention dictated complete abstinence from alcohol for these women. Portfolio 3A was defined as an intervention to eliminate smoking among all smokers. Finally, Portfolio 4A was an intervention that imposed both alcohol elimination and smoking cessation strategies.

Set B intervention portfolios combining prompt initiation of modern ART with *realistic reductions* in the prevalence of alcohol consumption and/or smoking were defined as follows. Portfolio 2B combined prompt initiation of modern ART with an intervention to reduce alcohol consumption of over 3 drinks per week to a 3 drink per week limit among those without an indication of HCV. While current guidelines classify alcohol consumption of over 7 drinks per week for women as "heavy drinking" beyond safe limits,[119] they refer to the general population and are not specific to those with HIV. Thus, to account for recent evidence suggesting that those with HIV are adversely impacted by alcohol at lower thresholds of consumption than the general population,[47] we considered women consuming over 3 drinks per week for this intervention.

For those with an indication of HCV at baseline, we based our intervention on a brief alcohol treatment program.[120] The intervention was conducted among patients reporting current drinking when they entered a Veterans Administration HCV clinic. Patients received brief alcohol counseling by trained HCV clinicians that included assessment of alcohol use, feedback regarding personal risk, education

regarding the interaction between alcohol and HCV on liver fibrosis and effects of alcohol on antiviral treatment, and advice regarding changes to consumption. Patients had a follow-up visit 4 to 8 weeks later with a psychiatric clinical nurse specialist at which they were provided cognitive-behavioral therapy techniques and motivational enhancement therapy for coping with cravings and preventing relapse. Among those who received the intervention, 36% achieved abstinence and 26% achieved >50% reduction in drinking levels after 8 months of follow-up.[120] Thus, we defined our intervention among those with an indication of HCV at baseline as a 0.36 probability of quitting alcohol, 0.26 probability of reducing consumption by 50%, and 0.38 probability of no change in alcohol consumption levels.

Portfolio 3B was defined as a behavioral smoking cessation intervention described by Hoffman et al[121] and tested in the WIHS. The intervention consisted of stage-based Transtheoretical model (TTM) tailored health communications on smoking cessation delivered via computer expert systems in waiting rooms at primary care clinics. The intervention was delivered at baseline, 3 months, and 6 months. Participants also received audiotapes relevant to their stage of change. At six months, 16.3% of participants had quit smoking. Thus, we defined our intervention among smokers as a 0.16 probability of quitting smoking.

The combination of the alcohol and smoking cessation interventions described above was defined as Portfolio 4B.

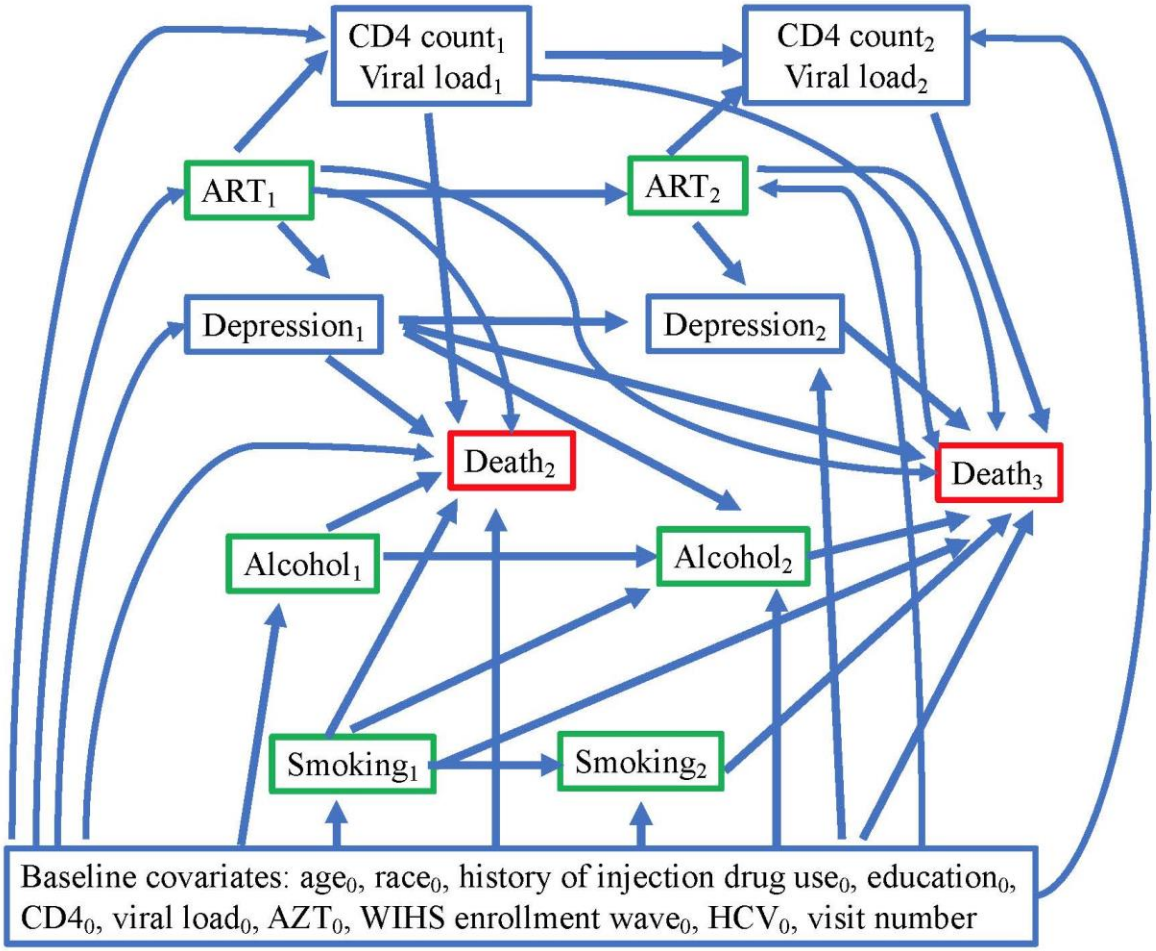
B5. Covariates of Interest

Demographic, laboratory, and risk factor data collected during WIHS semi-annual study visits were included in models as potential confounders. We selected covariates based on hypothesized relationships between the interventions and the outcome depicted by causal diagrams.[122] A simplified version of the causal diagram used to select time-fixed and time-varying confounders is shown in Figure 3.2. While the figure includes all exposures, it does not depict interactions between exposures. Time-fixed confounders included baseline age, race, level of education, history of injection drug use (IDU), prior exposure to 2-drug antiretroviral therapy, wave of WIHS enrollment, CD4 cell count, viral load, and smoking status. To simplify the representation of these relationships, the full set of time-fixed

confounders is shown in the figure though various elements from this set functioned as subsets of confounding variables for each exposure. Time-varying covariates included CD4 count, viral load, and depression.

We also considered a time-updated measure of visit number, interactions between time and each of the exposures (i.e., ART, alcohol, smoking) and interactions between exposures.

Figure 3.2. Causal Diagram of ART, Non-AIDS Risk Factors, and Mortality.*



* Relevant exposures are displayed in green, time-fixed and time-varying confounders are shown in blue, and the outcome is shown in red.

B6. Treatment Decision Design

The overall analytic goal of Aim 1 was to use rich, longitudinal data collected from the WIHS to obtain estimates of the potential reductions in the cumulative incidence of mortality that could be achieved by multiple interventions on non-AIDS risk factors. A defined set of criteria was used to determine eligibility for intervention among each WIHS participant at each study visit, with analyses structured in the format of a treatment decision design.[123] The treatment decision design is a generalization of the new-user design[113] which anchors a cohort's follow-up to times at which treatment decisions are made. While all WIHS participants meeting the inclusion criteria defined in Section B2 were eligible for inclusion in the analysis sample, only WIHS participants with a specific risk factor at a given visit were considered *eligible for the intervention* at that visit. We used semi-annual study visits to represent the timepoints at which intervention decisions were made. At each visit, women's risk factor information was used to determine eligibility for each intervention as described below.

ART: Because all women in the analysis sample were ART-naïve at baseline, they were all considered eligible for the intervention on prompt initiation of ART in the modern treatment era. Thus, all women were eligible to initiate ART by the first follow-up visit. Once women initiated ART, they were assumed to remain on ART and were no longer considered for ART initiation at later semi-annual visits.

Alcohol reduction: Both HCV status and alcohol consumption levels determined eligibility for the alcohol reduction intervention. Specifically, women with an indication of HCV at baseline were eligible for the alcohol intervention if they reported consumption of any alcohol at a given visit. Women without an indication of HCV at baseline were considered for the alcohol elimination intervention (Portfolio 2A) if they reported consumption of over 1 drink per week and for the alcohol reduction intervention (Portfolio 2B) if they reported consumption of over 3 drinks per week at a given visit. Women could receive this intervention multiple times.

Smoking cessation: Women reporting current smoking at a study visit were considered eligible at that visit as well as at all subsequent visits at which smoking was reported. Thus, women could receive this intervention multiple times.

B7. Notation and Parameters of Interest

Let uppercase letters denote random variables and lowercase letters denote potential realizations of those variables. Let i index 1, …, 1016 women in our study sample, j index 1, …, J completed visits of follow-up, and Y_{ij} represent an indicator of death for woman i at time j. The maximum number of follow-up visits is $J = 16$; as WIHS visits occur every 6 months, this corresponds to 8 years of follow-up since study baseline. We use A_{ij} to represent a binary indicator of antiretroviral therapy (ART) treatment, D_{ij} to represent categorical indicators for average drinking level per week since the last study visit and S_{ij} to represent a binary indicator of smoking. We use $\boldsymbol{Z}_{ij}$ to represent a vector of time-varying covariates for woman i measured at time j (i.e., CD4 count, detectable viral load, and depression) and $\boldsymbol{W}_{ij}$ to represent a vector of time-fixed demographic and clinical characteristics for woman i measured at baseline (i.e., $j = 0$). The vector $\boldsymbol{W}_{ij}$ includes baseline age, race, level of education, history of injection drug use (IDU), prior exposure to dual therapy zidovudine (AZT), wave of WIHS enrollment, CD4 count, viral load, and smoking status. Finally, we use C_{ij} to represent a binary indicator of being censored due to being lost to follow-up which we defined as having two consecutive missed WIHS visits. All women were administratively censored after a maximum of 8 years of completed follow-up due to sparse data at later time points. To accommodate sporadic missing data for time-varying covariates and exposures in the observed data, we used a last observation carried forward approach.

At each timepoint j, the temporal ordering of variables is: $\boldsymbol{W}_{ij}$; C_{ij}; Y_{ij}; A_{ij}, D_{ij}, S_{ij}; $\boldsymbol{Z}_{ij}$. The values of the time-fixed covariates $\boldsymbol{W}_{ij}$ remain constant for each woman throughout the study period so that for a given woman i, $\boldsymbol{w}_{i0} = \boldsymbol{w}_{i1} = \ldots = \boldsymbol{w}_{iJ}$. Among those who remained uncensored and alive by visit j (i.e., $C_{ij} = Y_{ij} = 0$), values of ART (i.e., A_{ij}), drinking level (i.e., D_{ij}), and smoking (i.e., S_{ij}) were measured based on self-report corresponding to the 6-month period between $j - 1$ and j.

It should be noted that while exposures to ART, drinking level, and smoking occurred between $j - 1$ and j, they were only documented at time j due to the nature of the WIHS as an interval cohort. It is unknown from the observed data whether, among those who die before a given follow-up visit, there was

a change in exposure status in the time period between the last follow-up visit at which they remained alive (and uncensored) and their death date. ART initiation dates are recorded as the first WIHS visit date at which individuals report having initiated ART, while death dates are recorded as actual dates of death. Therefore, we assume that those who had not initiated ART at the last visit at which they remained alive and uncensored did not do so before a death occurring before the next scheduled follow-up visit. Similarly, we assume drinking and smoking status did not change between the last visit at which women remained alive and uncensored and their death. Or rather, in both cases, we assume that any change in ART, drinking, or smoking between the last follow-up visit and a death occurring less than 6 months later was too short of an exposure period to notably affect the risk of death.

Among those who remain alive and uncensored at time j, measured values of time-varying covariates (i.e., $\boldsymbol{Z}_{ij}$) correspond to real time, or nearly real time, measures at j. Specifically, CD4 cell count and viral load were measured by laboratory tests conducted at the time of the WIHS visit. Depression was measured by a validated questionnaire[124] that asked women to report symptoms experienced during the last week. In the remaining text we suppress subscript i.

The cumulative incidence of mortality in the WIHS under the observed exposure history (i.e., no intervention on ART initiation or non-AIDS risk factors) at time j can be expressed as Equation 1:

$$F(j) = \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\bar{a}_j}\sum_{\boldsymbol{w}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_k=\bar{a}_k,\bar{D}_k=\bar{d}_k,\bar{S}_k=\bar{s}_k,\ \bar{\boldsymbol{Z}}_k=\bar{\boldsymbol{z}}_k,\boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\ \bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(S_m=s_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(D_m=d_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(A_m=a_m\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m)\times \end{bmatrix}\right\}$$

The cumulative incidence of mortality in the WIHS under universal, prompt initiation of ART in the modern treatment era can be expressed by adapting the above equation to set $A = 1$. Specifically, this indicator represents initiation of therapy by the time of the first follow-up visit (i.e., $A = 0$ for all women at $j = 0$ and $A = 1$ for all women at $j = 1$ and all visits thereafter). This is akin to an intent-to-treat analysis in which we assume that, once therapy is initiated, it is continued throughout the study period. It should be noted that women enrolled in the WIHS prior to 1 October 2001 and set to initiate ART by the

follow-up visit could not have initiated modern ART. This is based on our definition of modern ART initiation as initiation of ART on, or after, 1 October 2001 (the period after which tenofovir had been approved and became a common drug in combination ART regimens).[114] However, we used observed exposure information from those in the modern treatment era to extrapolate modern ART to this earlier time period. The adapted equation is written as Equation 2:

$$F^{a^g}(j) = \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\boldsymbol{w}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_{k-1}^{g}=\bar{a}_k^{g},\bar{D}_k=\bar{d}_k,\bar{S}_k=\bar{s}_k,\boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(S_m=s_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(D_m=d_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ 1\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m) \end{bmatrix}\right\}$$

where we use superscript g to convey that the value of A is being set according to the intervention plan.

Equation 2 can be further adapted to express the cumulative incidence of mortality in the WIHS under universal, prompt initiation of ART in the modern treatment era combined with various interventions on alcohol consumption and/or smoking. Because these interventions depend on the natural value of alcohol consumption and smoking, this requires use of the extended version of the parametric g-computation estimator.[104,105] This facilitates the expression of interventions that are implemented based on the values of alcohol and smoking that would be observed if intervention were discontinued immediately before measurement at time j. The incidence function under these additional interventions can be expressed with the extended parametric g-formula as Equation 3:

$$F^{(a,d,s)^g}(j)$$

$$= \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\boldsymbol{w}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_k^{g}=\bar{a}_k^{g},\bar{D}_k^{g}=\bar{d}_k^{g},\bar{S}_k^{g}=\bar{s}_k^{g},\ \bar{\boldsymbol{Z}}_k=\bar{\boldsymbol{z}}_k,\boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\ \bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\ \bar{S}_{m-1}^{g}=\bar{s}_{m-1}^{g},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{g}\left(S_m^{g}=s_m^{g}\middle|S_m^{*}=s_m^{*},\bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\ \bar{S}_m^{g}=\bar{s}_m^{g},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{obs}\left(S_m^{*}=s_m^{*}\middle|\bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\bar{S}_{m-1}^{g}=\bar{s}_{m-1}^{g},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{g}\left(D_m^{g}=d_m^{g}\middle|D_m^{*}=d_m^{*},\ \bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{obs}\left(D_m^{*}=d_m^{*}\middle|\bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ 1\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\ \bar{D}_{m-1}^{g}=\bar{d}_{m-1}^{g},\ \bar{S}_{m-1}^{g}=\bar{s}_{m-1}^{g},\ \bar{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m) \end{bmatrix}\right\}$$

We use superscript g to convey that the values of D and S (i.e., alcohol consumption and smoking) are being set according to the intervention plan while we use an asterisk to denote the values that would be observed if the intervention plan were discontinued immediately before measurement of

these variables. The value $\bar{a}_j^g$ is set to 1 across all $j \geq 1$ for each set of intervention portfolios while the values of $\bar{d}_j^g$ and $\bar{s}_j^g$ are set according to the plans described in Section B4.

B8. Implementation of Parametric G-computation

We used the parametric g-computation estimator[67–69,71,72,78,99,105] to estimate the equations in Section B7. Namely, we estimated the 8-year cumulative incidence of all-cause mortality under the natural course, the 8-year cumulative incidence of all-cause mortality under prompt initiation of modern ART, and the 8-year cumulative incidences of all-cause mortality under each set of intervention portfolios. Furthermore, we estimated the 8-year risk differences and ratios contrasting each intervention portfolio to an intervention solely on prompt initiation of modern ART. The steps to using the g-computation estimator to parametrically estimate the g-computation formula have been previously described.[68,69,71] Briefly, we implemented g-computation in our setting as follows.

Steps 1 – 3: First, we fit pooled person-visit logistic and linear models to the observed data across all visits $j \geq 1$ to estimate the conditional probability or density of each exposure, time-varying covariate, and outcome. Specifically, we fit models for ART initiation, alcohol consumption, smoking, CD4 cell count, viral load, depression, and death. Second, we estimated conditional probabilities or densities from these models. Third, we drew a large Monte Carlo sample of $N = 100{,}000$ women at baseline (i.e., $j = 0$) with replacement.

Step 4: We used the estimated conditional probabilities or densities and observed values of exposures and covariates at the baseline visit to simulate the values of exposures, covariates, and outcomes in the temporal order they would have occurred at $j = 1$ by drawing values from a Bernoulli distribution in the case of binary variables and multinomial distributions in the case of categorical variables. We then used the estimated conditional probabilities or densities and simulated values of exposures, covariates, and outcomes at $j = 1$ to simulate the values of these variables at $j = 2$, continuing this process for a maximum of 16 follow-up visits to ultimately generate a dataset representing the

outcome distribution under the natural course (Equation 1 in Section B7). By natural course, we mean no intervention (i.e., no manipulation of treatment status).

To achieve this, we first fit intercept only models to the pooled person-visit data for each time-varying covariate, exposure, and outcome with indicator variables for visit number as the only predictor. We then compared the distribution of each variable at each visit in the Monte Carlo dataset with the distribution of each variable at each visit in the original WIHS dataset. Further complexity was then incorporated into each of the models by adding relevant covariates (e.g., hypothesized confounders, predictors of the dependent variable) and interactions and adapting functional forms. After each specification of the set of models, conditional probabilities or densities were estimated and saved from each model, a Monte Carlo sample of size 100,000 was drawn from the baseline data with replacement, follow-up data were generated at each follow-up visit from the saved conditional probabilities or densities, and the distribution of variables in the Monte Carlo dataset were compared with the distribution in the WIHS dataset using figures and summary statistics (e.g., frequencies, mean/median/standard deviation). We also assessed interaction terms using Likelihood Ratio Tests for interaction.

We estimated the cumulative incidence of mortality at each time point using the complement of the Kaplan-Meier[125] (product-limit) estimator of the survival function. The estimator can be expressed as:

$$\hat{S}(j) = \prod_{k=1}^{j} \left(1 - \frac{d_k}{n_k}\right)$$

where d_k denotes the number of deaths at discrete time k, n_k denotes the number at risk at discrete time k, and $\hat{S}(j)$ denotes the estimated survival at time j.

If parametric models are correctly specified, the cumulative incidence of mortality estimated in the Monte Carlo dataset under the natural course (i.e., no intervention on exposures) should be equivalent to the cumulative incidence of mortality observed in the WIHS dataset. Similarly, the distribution of each exposure and time-varying covariate in the Monte Carlo sample should be equivalent to the distributions

in the observed WIHS data. Discrepancies can arise from improper model specification as well as finite sample bias resulting from sparse data.

Thus, to improve model fit, we repeated the above Steps 1 – 4 using flexible specifications of continuous and categorical variables, interaction terms, and various sets of covariates for each model until we 1) had included all hypothesized confounders relevant to each exposure, and 2) were able to replicate distributions observed in the WIHS sample in the Monte Carlo sample. The variables and models that were selected for final analyses are described in Appendix A of the Supplemental Text.

Step 5: Once we were able to replicate the empirical distribution of covariates and outcome in the Monte Carlo sample, we used the finalized models to simulate outcomes under the interventions of interest. We repeated Steps 1 – 4, altering exposure values according to the defined interventions and simulated follow-up data using the saved conditional probabilities or densities.

Step 6: We estimated the cumulative risk of all-cause mortality at each follow-up visit in each dataset using the complement of the Kaplan-Meier[125] (product-limit) estimator of the survival function. We censored all women at a maximum of 16 follow-up visits (i.e., 8 years of follow-up) due to sparse data at later time points. We output the cumulative risks at each timepoint and concatenated datasets. To quantify the impact of scaling up interventions on non-AIDS risk factors, we contrasted the estimated risk of all-cause mortality under each intervention portfolio with the estimated risk under intervening on prompt initiation of ART in the modern treatment era alone using risk differences and risk ratios.

Step 7: We repeated the above steps on 1000 nonparametric bootstrap[126] samples to obtain standard errors calculated as the standard deviation of point estimates across bootstraps. We used standard errors to calculate 95% confidence intervals.

C. Aim 2: Development and Implementation of Two-Stage G-computation Estimators

The overall goal of Aim 2 was to develop and assess two-stage g-computation estimators of treatment and intervention effects in a given HIV cohort when exposure information is missing for a subset of the sample.

C1. Motivating Example

The primary example for our development and assessment of two-stage g-computation estimators focused on a simplified time-fixed scenario. In the first part of the scenario, interest was in obtaining unbiased estimates of the covariate vector $\boldsymbol{W}$ standardized risk of an outcome Y in an HIV cohort under scenarios in which the entire cohort did not receive versus did receive a harmful treatment A. This parameter is known as the average treatment effect.[46,49,88,103] In the second part of the scenario, interest was in obtaining unbiased estimates of the covariate vector $\boldsymbol{W}$ standardized risk of an outcome Y in the HIV cohort under a scenario in which an intervention removed treatment A versus treatment practices remained the same (i.e., the natural course or status quo). This parameter is known as the average intervention effect.[46,49,51,88,103] More specifically, it represents a special case of an generalized intervention effect in which the intervention is 100% efficacious.[51]

While it is straightforward to implement the parametric g-computation estimator to obtain estimates of these parameters when data are complete, this is no longer the case when faced with partially missing exposure information. If estimates are desired for the total study population but treatment status is only observed on a subset (i.e., where $R = 1$) complications arise, especially if there are other variables associated with the outcome that additionally influence whether or not treatment status is observed; in other words, if there exists a variable Z associated with both R and Y.

To concretize this scenario, we considered a hypothetical study in which the objectives were to estimate the average treatment and intervention effects[46,49,51,88,103] of opioid prescription duration in the Women's Interagency HIV Study (WIHS) on the 12-month risk of hospitalization/emergency room visit or death from any cause. Opioids are prescribed at a higher rate among those with HIV[127] whom already experience a higher prevalence of substance abuse and addiction.[55] Initial opioid prescribing practices influence long-term opioid use[128] with duration of initial prescription being the strongest predictor of opioid use disorder and overdose.[129] Thus, it is reasonable that investigators might be interested in outcomes under various opioid duration prescribing strategies among those with HIV who receive opioid prescriptions. Therefore, we considered a case in which our goal was to estimate the average treatment

effect of shortening versus lengthening the duration of initial opioid prescriptions among all women and the average intervention effect of shortening the duration of initial prescriptions versus status quo prescribing practices.

C2. Study Data

We based our example on a hypothetical cohort simulated to represent HIV-positive women enrolled in the WIHS (described in Section B1). We created our dataset such that the full sample size was equal to the number of HIV-positive women enrolled in waves 2-4 of WIHS (i.e., $N = 1623$) and the incidence of hospitalization/emergency room visit or death over 12 months was similar to that observed in a sample of outcome data from the most recent wave of enrollees and which has been reported in previous HIV studies[130] (i.e., approximately 30%).

While information on opioid use has been collected on WIHS participants, this information has primarily been based on self-reported use of illicit or recreational drugs. Because our objective was to estimate the effect of a clinical intervention in which guidelines dictate that physicians shorten the period over which initial opioid prescriptions are provided, we considered a scenario in which it was possible to supplement the existing WIHS cohort data with linkages to electronic medical record databases to obtain detailed dosage and duration information for individuals receiving prescriptions. Theoretically, this could be done for the WIHS, but data linkages are expensive and often cannot be performed for the entire study sample. Thus, it is more likely that only a sample of linkages would be performed. Specifically, we considered a case in which it was possible to link prescription information for a nonrandom subset of 30% of the simulated cohort, particularly members receiving HIV care at 10 clinics proximal to the WIHS sites.

A causal diagram[122] for our research scenario is shown in Chapter 5, Figure 5.1. Receipt of an initial long duration opioid prescription at baseline (i.e., A) increases the risk of emergency room/hospital visit or death by 12 months (i.e., Y). History of substance use (i.e., W) represents one potential confounder increasing the probability of receiving an initial long duration prescription and the risk of emergency room/hospital visit or death. Not receiving regular HIV care (i.e., Z), while not a confounder

and therefore often omitted from the causal diagram, decreases the probability of having observed exposure information (i.e., $R = 1$) from linked electronic medical records and also increases the risk of hospitalization or death (i.e., Y).

According to this scenario, we generated n = 1 to 1623 independent and identically distributed sets $\{W, A, Z, R, Y^{a=0}, Y^{a=1}, Y\}_n$ representing the N = 1623 members in the full population of interest. A represents exposure to long (i.e., $A = 1$) versus short (i.e., $A = 0$) duration of opioid prescription; W represents history of substance use (history: $W = 1$; no history: $W = 0$); Z represents lack of engagement in regular HIV care (lack of engagement: $Z = 1$; engagement: $Z = 0$); R represents observed exposure information (i.e., selection into the sample of complete cases, $R = 1$); $Y^{a=0}$ represents the potential outcome for emergency room visit or death under short term opioid exposure; $Y^{a=1}$ represents the potential outcome for emergency room visit or death under long term opioid exposure; and Y represents the factual outcome for emergency room visit or death given the exposure received.

<u>For each record, covariate values were simulated in SAS 9.4 in the following order:</u>

W drawn from a Bernoulli distribution with a marginal prevalence of 0.44

i.e., w = rand("BERNOULLI",0.44);

Z drawn from a Bernoulli distribution with a marginal prevalence of 0.4

z = rand("BERNOULLI",0.4);

R drawn from a Bernoulli distribution as a function of Z and with a marginal prevalence of 0.3

r = rand("Bernoulli",1/(1+exp(-(-log(1/0.26-1)-log(8)*z+log(8)*0.4))));

$Y^{a=0}$ drawn from a Bernoulli distribution as a function of $W$and Z and with a marginal incidence of 0.19

y0 = rand("BERNOULLI",1/(1+exp(-(-log(1/.19-1)-log(1.5)*0.44+log(1.5)*w1-log(1.5)*0.4+log(1.5)*z))));

$Y^{a=1}$ drawn from a Bernoulli distribution as a function of $W$and Z and with a marginal incidence of 0.37

y1 = rand("BERNOULLI",1/(1+exp(-(-log(1/.35-1)-log(1.5)*0.44+log(1.5)*w1-log(10)*0.4+log(10)*z))));

A drawn from a Bernoulli distribution as a function of W and with a marginal prevalence of 0.6

a = rand("BERNOULLI",1/(1+exp(-(-log(1/.6-1)-log(2)*0.44+log(2)*w1))));

Y set according to realized value a

if a then y=y1; else y=y0;

The potential outcomes $Y^{a=0}$ and $Y^{a=1}$ were generated as a way of directly setting the truth by which to evaluate the performance of estimators. Setting observed Y to Y^1 or Y^0 according to observed a is an application of the consistency assumption (discussed in Chapter 3, Section A2). The true risks and risk differences were determined using the potential outcomes $Y^{a=1}$, $Y^{a=0}$, and Y from the full cohort prior to imposing missing exposure information.

C3. Estimators

Let A denote a binary time-fixed exposure, $\boldsymbol{W}$ denote a vector of confounders, Y denote an outcome, uppercase letters denote random variables, and lowercase letters denote possible values of variables or constants. To simplify notation, we assume $\boldsymbol{W}$ to be a finite dimension vector of categorical covariates. When the vector includes continuous covariates equations are written as integrals. The expected value of the potential outcome Y^a under a time-fixed exposure plan a can be written as,

$$(\text{C3.1}) \quad E[Y^a] = \sum_{\boldsymbol{w}} E(Y|A = a, \boldsymbol{W} = \boldsymbol{w})P(\boldsymbol{W} = \boldsymbol{w})$$

where we define $E[Y^{a=1}]$ as the expected outcome distribution under a plan in which the full study sample receives exposure and $E[Y^{a=0}]$ as the expected outcome distribution under a plan in which the full study sample does not receive exposure. The expected value of the outcome under the natural course (i.e., observed exposure plan) can be written as,

$$(\text{C3.2}) \quad E[Y] = \sum_{\boldsymbol{w},a} E(Y|A = a, \boldsymbol{W} = \boldsymbol{w})P(A = a|\boldsymbol{W} = \boldsymbol{w})P(\boldsymbol{W} = \boldsymbol{w}).$$

The process of estimating the above g-formula expressions by g-computation has been described in detail in the literature[68,69,71,98] and was reviewed in the application in Aim 1 (Chapter 3, Section B8). As a brief review as it applies to this simplified time-fixed setting, exposure and covariate conditional probabilities or densities of the outcome are estimated from models fit to the full study sample, a large Monte Carlo sample is drawn from the original study population at baseline, exposure values for each individual in the Monte Carlo sample are set according to the exposure plan, and outcomes are simulated in the Monte Carlo sample using the set exposure value a, realized covariate values $\boldsymbol{w}$, and conditional probabilities that were estimated in the first step. To obtain the outcome distribution in the total study population, outcomes are summed across the population's distribution of $\boldsymbol{w}$. It is standard practice to estimate the outcome distribution under the natural course with g-computation regardless of whether it is a parameter of interest; comparison of the observed natural course to the g-computation simulated natural course is used as a check on model specification prior to using the models to simulate outcome distributions under altered exposure plans.

When exposure plan information for some participants is missing, missing data prohibit the use of the above approach. Instead, under an assumption that patients with (i.e., $R = 1$) and without (i.e., $R = 0$) complete exposure information are exchangeable within levels of confounders W, investigators often use a complete case analysis. Under this assumption, the g-formula equations for the potential outcome under exposure plan a and under the natural course can be written as:

$$\text{(C3.3)} \quad E[Y^a] = \sum_{\boldsymbol{w}} E(Y|A = a,\ \boldsymbol{W} = \boldsymbol{w}, R = 1)P(\boldsymbol{W} = \boldsymbol{w}, R = 1)$$

$$\text{(C3.4)} \quad E[Y] = \sum_{\boldsymbol{w},a} E(Y|A = a,\ \boldsymbol{W} = \boldsymbol{w}, R = 1)P(A = a|\boldsymbol{W} = \boldsymbol{w}, R = 1)P(\boldsymbol{W} = \boldsymbol{w}, R = 1).$$

The estimation of conditional outcome probabilities or densities is restricted to the subsample with observed exposure (i.e., complete information), and the outcome distribution is summed over the distribution of $\boldsymbol{w}$ among those with complete information. However, $E[Y^a|W = w]$ will only be

approximated by $E[Y^a|W = w, R = 1]$ if exposure information is missing completely at random or missing at random given the vector of confounders $\boldsymbol{W}$.

Two-stage g-computation estimator of average treatment effect $E[Y^{a=0} - Y^{a=1}]$

We proposed a g-computation extrapolation estimator of the average treatment effect which combines partial information on the full study sample with complete information on the subset of participants with observed exposure. We rewrite the g-formula to 1) include $\boldsymbol{Z}$ in our outcome model, where $\boldsymbol{Z}$ is a vector of covariates associated with R and Y that allows us to meet the assumption that presence of exposure information is independent of Y given $\boldsymbol{W}$ and $\boldsymbol{Z}$ (i.e., $E[Y^a|a, \boldsymbol{w}, \boldsymbol{z}] = E[Y^a|a, \boldsymbol{w}, \boldsymbol{z}, R = 1]$), and 2) sum outcomes over the joint distribution of $\boldsymbol{W}$ and $\boldsymbol{Z}$ instead of only $\boldsymbol{W}$ in the subset with complete exposure information. The g-computation extrapolation estimator is written as:

(C3.5) $\quad \hat{E}_{ext}[Y^0 - Y^1] =$

$$\sum_{\boldsymbol{w},\boldsymbol{z}} \hat{P}(Y = 1|A = 0, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)\hat{P}(\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z})$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z}} \hat{P}(Y = 1|A = 1, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)\hat{P}(\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}) .$$

Stage 1 consists of estimating $\hat{P}(Y = 1|A = a, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)$ from a logistic model fit to complete cases; stage 2 consists of (a) simulating outcomes under plans $a = 0$ and $a = 1$ among Monte Carlo resamples taken from the *full* population (i.e., not only sampling complete cases) and then (b) summing over the full population's joint distribution of $\boldsymbol{w}$ and $\boldsymbol{z}$. We obtain the average treatment effect by taking the difference $\hat{E}[Y^0 - Y^1]$.

Two-stage g-computation estimators of average intervention effect $E[Y^{a=0} - Y]$

The risk under the natural course is estimated by additionally summing over the full population's distribution of observed exposure. Because the intervention effect contrasts the outcome distribution under an altered exposure distribution to the outcome distribution under the observed exposure distribution, an additional step is needed when exposure status is missing for a subset of the sample and the outcome distribution differs among the subset with observed information versus the full sample.

Similarly, this is needed for interventions that depend on the natural value of exposure[104] such as for the effect estimated by Lesko et al.[64] in which intervening on depression depended on the natural value (i.e., observed level) of depressive symptoms.

We proposed two-stage inverse probability-weighted (IPW) and exposure imputation g-computation estimators of the average intervention effect. As with the extrapolation g-computation estimator, IPW and imputation g-computation include $\boldsymbol{Z}$ in the outcome model as well as the additional model specified for missing data.

IPW g-computation estimator:

(C3.6) $\hat{E}_{IPW}[Y^0 - Y] =$

$$\sum_{\boldsymbol{w},\boldsymbol{z}}\left[\frac{\hat{P}(Y=1|A=0,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z}|R=1)}{\hat{P}(R=1|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})}\right]$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z},a}\left[\frac{\hat{P}(Y=1|A=a,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(A=a|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z}|R=1)}{\hat{P}(R=1|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})}\right].$$

Stage 1 consists of estimating the probability of being a complete case (i.e., $\hat{P}(R=1\,|\,\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$)[131] in the *full* population in addition to estimating $\hat{P}(Y=1|A=a,\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},R=1)$ in the complete cases as in the extrapolation approach. In stage 2, we (a) simulate outcomes setting $a=0$ and $A=a$ under the observed exposure value among Monte Carlo resamples taken from the *complete cases*, (b) weight simulated outcomes by the inverse probability of being a complete case, (c) and sum over the joint distribution of $\boldsymbol{W}$ and $\boldsymbol{Z}$ in the *complete cases.*

Exposure imputation g-computation estimator:

(C3.7) $\hat{E}_{impute}[Y^0 - Y] =$

$$\sum_{\boldsymbol{w},\boldsymbol{z}}\hat{P}(Y=1|A=0,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z},a}\hat{P}(Y=1|A=a,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(A^*=a|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$$

where a denotes the realized value of observed exposure or imputed exposure (where missing). A separate model is specified in stage 1 to estimate the conditional probability of exposure among those with complete information, (i.e., $\hat{P}(A = 1|\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, Y = y,\ R = 1)$); in stage 2, (a) values of exposure are imputed for each woman with missing exposure information in the Monte Carlo resamples taken from the *full* population using the conditional probabilities estimated in stage 1, and (b) outcomes are simulated under the *full* joint distribution of imputed or observed a^*, $\boldsymbol{w}$, and $\boldsymbol{z}$. Because no values for a are imputed when a is set to 0, the first part of the equation for Y^0 in equation C3.7 is identical to that of the extrapolation approach in C3.5.

C4. Statistical Analyses

We used proposed two-stage g-computation estimators as well as the naïve, complete case g-computation estimator to estimate parameters of interest in our simulated cohort of 1623 HIV-positive women. In our generated cohort, we set exposure to missing where $R = 0$. We fit the following logistic model for the outcome among complete cases to estimate the conditional probability of hospital/emergency room visit or death at 12-months $\hat{P}(Y = 1|A = a,\ W = w, Z = z,\ R = 1) = expit(\alpha_0 + \alpha_1 a + \alpha_2 w + \alpha_3 z)$, where $expit\ = \exp(\alpha)\ /[1 + \exp(\alpha)]$. We used a Firth correction[132] to improve model convergence. We also fit the following logistic model for the probability of being a complete case (i.e., having complete prescription duration information) $\hat{P}(R = 1|W = w, Z = z) = expit(\tau_0 + \tau_1 w + \tau_2 z)$ and took the inverse of the probability for the complete case weight.[131] Finally, we fit the following logistic model for the conditional probability of receiving a long duration opioid prescription for >7 days $\hat{P}(A = 1|W = w, Z = z) = expit(\delta_0 + \delta_1 w + \delta_2 z)$.

After saving the estimated conditional probabilities, we took a Monte Carlo sample of size 10,000 of the full cohort. In the Monte Carlo sample, we used the two-stage extrapolation, inverse probability weighted, and exposure imputation g-computation estimators to estimate the 12-month probability of hospital/emergency room visit or death if all women were to receive short duration opioid prescriptions for ≤7 days (i.e., $\hat{E}[Y^{a=0}]$ and if all women were to receive long duration opioid prescriptions for >7

days (i.e., $\hat{E}[Y^{a=1}]$). We contrasted these two estimates to obtain risk differences for the average treatment effect. We additionally used the two-stage inverse probability weighted and exposure imputation g-computation estimators to estimate the risk if prescription practices were to remain unchanged (i.e., $\hat{E}[Y]$) and the average intervention effect if all women received short duration opioid prescriptions versus the status quo. We estimated the standard error based on the standard deviation of 200 nonparametric bootstrap resamples to obtain 95% confidence intervals.[126] We accounted for error in the estimation of inverse probability weights and exposure imputation by performing these steps within the bootstrap samples.

To assess estimator performance, we repeated analyses in 10,000 simulation trials to compute bias, standard error, root mean squared error, and confidence limit coverage. We defined bias as the difference between the value of the parameters of interest estimated with each approach and the true value of the parameters of interest obtained from the potential outcomes prior to imposing missing data. Standard errors of the risks, risk differences, and bias were calculated based on the standard deviation of the point estimates of the values for these parameters across all trials. The root mean squared error was calculated as the square root of the estimated sum of the squared bias and variance. 95% confidence limit coverage was defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation experiments.

To further assess the operating characteristics of proposed estimators, we repeated 10,000 simulation trials under varying sample conditions and analogously estimated bias, standard error, root mean squared error, and confidence limit coverage in these settings. We varied the percentage of complete exposure data in 5% increments from 10-90% to assess the robustness of estimators to the proportion of missing data as well as to determine the point at which the naïve g-computation estimator fit to complete cases may be minimally biased. We also repeated simulation trials excluding the covariate Z from the outcome model to test our hypothesis that all two-stage estimators would require the inclusion of this covariate in the outcome model in addition to standard confounders. We simulated data under the null treatment and intervention effects to confirm estimators were valid (i.e., we simulated the risk under

shortened, lengthened, and natural course opioid durations to be equal so that the risk differences for the average treatment and intervention effects were 0). In another scenario, we simulated missing data independent of Z so that data would be missing completely at random. This was to test our hypothesis that the standard g-computation estimator would be unbiased in this restrictive setting. We increased the strength of modification of the treatment and intervention effects by Z to assess whether there would be a greater improvement in bias with two-stage g-computation estimators over the standard g-computation estimator in these cases. Finally, in small samples sizes, large proportions of missing data are likely to create issues of data sparsity. Because our base scenario included a cohort of 1623 women and small proportions of completely observed exposure, we wanted to assess whether any potential bias in estimates may be due to finite samples. Thus, we repeated trials in an increased cohort size of 7500 women to assess the operating characteristics of estimators in settings without possible violations in positivity.

CHAPTER 4: ALL-CAUSE MORTALITY UNDER MODELED INTERVENTIONS ON ANTIRETROVIRAL THERAPY, ALCOHOL, AND SMOKING AMONG HIV-POSITIVE WOMEN IN THE UNITED STATES, 1998 – 2017

A. Overview

People with HIV experience a higher prevalence of multimorbidity, accelerated aging, and decreased survival compared to the general US population despite two decades of effective antiretroviral therapy (ART). Yet, the improvement in survival that may be achieved by targeting non-AIDS risk factors among those who initiated antiretroviral therapy (ART) remains unclear.

Leveraging observational data from 1016 HIV-positive women participating in the Women's Interagency HIV Study between 1998 and 2017, we estimated the mortality risk under modeled interventions targeting alcohol consumption and smoking with ART initiation. We used parametric g-computation to estimate effects of combining ART with interventions eliminating or reducing the prevalence of alcohol and/or smoking compared to an intervention solely on universal initiation of ART in the modern treatment era.

Among 1016 women at baseline, 59% reported smoking and 23% reported consuming over 3 drinks per week. The observed 8-year risk of mortality was 22.5% compared to an estimated 10.4% (95% CI: 6.3, 14.5) risk under universal initiation of ART in the modern treatment era. The 8-year risk differences contrasting universal initiation of modern ART and elimination of non-AIDS risk factors, with intervening on ART alone, were -0.5% (95% CI: -1.2, 0.3) with near elimination of alcohol, -1.8% (95% CI: -3.6, 0) with 100% smoking cessation, and -2.0% (95% CI: -3.9, -0.2) with both. Under interventions where the probability of abstaining from alcohol or quitting smoking was less than 1, reflecting the reality that interventions rarely eliminate exposures entirely, reductions in the 8-year risks of mortality were less pronounced but still suggested improvements in survival compared to intervening on ART alone.

While modern ART has transformed the prognosis of HIV to a manageable chronic condition, smoking and alcohol remain significant contributors to early mortality. Interventions that target these risk factors, particularly smoking, may further reduce the risk of mortality, but more efficacious interventions are needed.

B. Introduction

Non-AIDS causes of morbidity and mortality disproportionately affect those with HIV, despite two decades of effective antiretroviral therapy (ART).[1–4] Compared to the general US population, people with HIV experience a higher prevalence of multimorbidity,[16] accelerated aging,[17–19] and decreased life expectancy, dying from non-AIDS causes 8 to 9 years earlier than those without HIV.[18,20,21] Thus, further improvements in the health status and survival of those with HIV will require greater attention to non-AIDS risk factors.

Cigarette smoking, heavy alcohol consumption, and chronic inflammation are highly prevalent among those with HIV and may contribute to a large proportion of excess mortality.[23,25–28,30,133,134] Studies suggest a synergistic relationship between cigarette smoking and HIV infection[15,41–44,46] as well as an increased risk of physiologic injury and mortality at lower thresholds of alcohol consumption among those with HIV.[47] Yet, the improvement in long-term outcomes that may be achieved by targeting non-AIDS risk factors along with prompt initiation of modern ART regimens remains unclear.

Here, we estimate the 8-year risks of all-cause mortality in the Women's Interagency HIV Study (WIHS) under proposed interventions that combine 1) prompt initiation of ART in the modern treatment era with 2) alcohol reduction, 3) smoking cessation, and 4) both alcohol reduction and smoking cessation. We quantify the improvement in survival that may be achieved by each intervention combination compared to intervention on prompt ART initiation alone.

C. Methods

C1. Study Sample

The Women's Interagency HIV Study (WIHS) is an ongoing, multisite prospective cohort study of HIV-positive and HIV-negative women in the United States.[89–91] The WIHS includes 10 consortia of

contributing sites (Bronx/Manhattan, NY; Brooklyn, NY; Chicago, IL; Los Angeles/Southern CA/Hawaii; San Francisco/Bay Area, CA; Washington, DC; Atlanta, GA; Birmingham, AL/Jackson, MS; Chapel Hill, NC; Miami, FL). Women were enrolled during 4 recruitment waves (1994-1995, 2001-2002, 2011-2012, and 2013-2015) and have since been followed every 6 months for documentation of sociodemographic information, risk behaviors, lab measurements, medications, physical exams, and clinical events. Since 1994, the WIHS has enrolled 4982 women, 3703 of whom were HIV-positive at baseline or seroconverted over follow-up. Institutional review boards at each WIHS site approved the WIHS and all women provided written informed consent for participation.

Eligibility for analysis was restricted to HIV-positive women with a study visit on, or after, 1 April 1998, and who had not previously initiated combination ART (N=1033). We excluded 2% with missing baseline information on depression (n=17), alcohol consumption (n=9), and smoking status (n=8). In the remaining sample of 1016 HIV-positive women, we imputed 5 (0.5%) missing baseline CD4 counts and 18 (1.8%) missing baseline viral load counts by simulation from log normal distributions with mean and standard deviation equal to that of the observed WIHS sample at baseline. The final sample for analysis included N=1016 HIV-positive, combination ART-naïve, women. This analysis was deemed exempt from review by the Office of Human Research Ethics at the University of North Carolina-Chapel Hill.

C2. Exposures and Outcome

Information on ART, alcohol consumption, and smoking has been collected on WIHS participants via in-depth interviews at baseline and follow-up visits. Combination ART is determined using a definition based on Department of Health and Human Services/Kaiser Panel guidelines.[114] Participants are considered to be on combination ART if they report use of 3+ antiretroviral drugs, one of which is a protease inhibitor (PI), a non-nucleoside reverse transcriptase inhibitor (NNRTI), an integrase inhibitor (II), or an entry inhibitor (EI). Here, we defined modern ART as combination ART initiated on/after 1 October 2001. This date coincides with the period after which tenofovir had been approved and

became a common drug in combination ART regimens.[114] During interviews, participants also reported the number of drinks consumed on average per week and smoking status since the last visit.

The outcome of interest was death due to any cause. The WIHS confirms deaths for all women who have been enrolled in the cohort through regular queries of state death certificates and the National Death Index Plus.

C3. Intervention Portfolios

Our primary objective was to estimate the risk[115,116] of all cause-mortality in the WIHS under hypothetical intervention portfolios (i.e., intervention combinations) targeting ART, alcohol, and smoking. In our base scenario (Portfolio 1), we considered an intervention that solely targeted prompt initiation of modern ART, where "prompt" initiation was defined as initiation of ART by the first follow-up visit (i.e., 6 months after the baseline visit). In our three other scenarios, we considered portfolios combining prompt initiation of modern ART with an intervention on: alcohol (Portfolio 2), smoking (Portfolio 3), and both alcohol and smoking (Portfolio 4). Specifically, we defined interventions on non-AIDS risk factors in two ways. Set A intervention portfolios combined ART with *near elimination* of alcohol consumption and/or smoking whereas Set B intervention portfolios combined ART with *realistic reductions* in the prevalence of alcohol consumption and/or smoking based on the expected efficacy of real-world interventions.

Set A intervention portfolios were defined as follows. Portfolio 2A was an intervention to eliminate alcohol consumption of over 1 drink per week among women without hepatitis C virus and eliminate all alcohol consumption among women with hepatitis C virus (i.e., positive antibody and positive or missing RNA status at baseline). We chose a limit of 1 drink per week instead of complete abstinence due to potential protective effects[117] of alcohol at low levels but did not define the intervention to increase drinking among those who were already abstainers. Because alcohol is particularly harmful for those with hepatitis C infection,[118] the intervention dictated complete abstinence from alcohol for these women. Portfolio 3A was defined as an intervention to eliminate smoking among all smokers. Finally,

Portfolio 4A was an intervention that imposed both near alcohol elimination and smoking cessation strategies.

Set B intervention portfolios to reduce, rather than eliminate, non-AIDS risk factors were specified to facilitate estimation of generalized intervention contrasts[51] and were defined as follows. Portfolio 2B combined prompt initiation of modern ART with an intervention to reduce alcohol consumption of over 3 drinks per week to a 3 drink per week limit among those without hepatitis C virus. This limit was chosen to account for recent evidence suggesting that those with HIV are adversely impacted by alcohol at lower thresholds of consumption than the general population.[47] Consequently, current guidelines[119] that classify alcohol consumption of over 7 drinks per week for women in the general population as "heavy drinking" beyond safe limits are likely too lenient for those with HIV. For those with hepatitis C virus at baseline, we based our intervention on a brief alcohol treatment program that resulted in alcohol abstinence among 36% of participants and >50% reduction in drinking levels among 26% of participants.[120] Portfolio 3B was defined as a behavioral smoking cessation intervention tested by Hoffman et al.[121] in which smokers had a 0.16 probability of quitting by 6 months. The combination of the alcohol and smoking cessation interventions described above was defined as Portfolio 4B.

C4. Analyses

We used the parametric g-computation algorithm,[67–69,71,73,75,98,104,105] a generalization of standardization, to estimate the mortality risk under each intervention portfolio. Under a sufficient set of conditions,[107–109,111,135] this approach re-expresses the mortality distribution under the observed treatments/exposures as the mortality distribution that would be observed under the intervention portfolios. Technical details are provided in the supplement. Briefly, we specified pooled logistic and linear person-visit models to estimate the probability or density of each exposure (i.e., antiretroviral therapy, level of alcohol consumption, smoking status), time-varying covariate (i.e., CD4 count, detectable viral load, depression), and outcome (i.e., mortality) conditional on observed values of covariates at each time period. Time-fixed covariates included baseline age, race, CD4 count, viral load,

education, history of injection drug use, WIHS recruitment wave, and prior exposure to 2-drug antiretroviral therapy.

For all analyses, we estimated the mortality risk using the complement of the Kaplan-Meier[125] survival function and administratively censored women at 8 years of follow-up due to sparse data at later time points. We used risk differences and ratios to quantify the impact of jointly targeting non-AIDS risk factors by comparing the 8-year risks of mortality under each portfolio to the 8-year risk of mortality with an intervention on prompt initiation of modern ART alone. We estimated 95% confidence intervals from the standard deviation of 1000 nonparametric bootstrap resamples.[126]

D. Results

Among 1016 HIV-positive women in the WIHS, the median age at study entry was 38 [Interquartile Range (IQR): 32, 45], median CD4 cells/mm^3 was 452 [IQR: 273, 655], and median $\log_{10}$ viral load copies/mL was 3.6 [IQR: 2.8, 4.4] (Table 4.1). Almost 70% of the cohort was Black, half reported depressive symptoms, 33% had a history of injection drug use, and 33% had evidence of hepatitis C virus. The smoking prevalence was 59% with an additional 15% of women reporting former smoking. Nearly half the study population abstained from alcohol consumption while 23% reported consuming over 3 drinks on average per week, the majority of whom were heavy drinkers (i.e., consuming over 7 drinks per week). The 8-year risk of all-cause mortality under the observed initiation of ART and no additional intervention on non-AIDS risk factors was 22.5% (Figure 4.1).

The risk of all-cause mortality was 10.4% (95% CI: 6.3, 14.5) under an intervention on prompt initiation of modern ART alone. Figure 4.2 displays the 8-year risk functions of all-cause mortality under interventions that combined prompt initiation of modern ART with near elimination (Panel A) or reduction (Panel B) of alcohol and smoking; numerical results are presented in Tables 2 and 3 respectively. Overall, the risk of all-cause mortality was lower under each intervention combining prompt initiation of modern ART with interventions on non-AIDS risk factors than under prompt initiation of modern ART alone. In both elimination and reduction strategies, interventions that targeted smoking were more beneficial than those that targeted alcohol consumption.

The risk of all-cause mortality by 8 years was 9.9% (95% CI: 5.9, 14.0) with near elimination of alcohol consumption, 8.6% (95% CI: 5.1, 12.1) with 100% smoking cessation, and 8.4% (95% CI: 4.3, 11.9) with near elimination of both alcohol and smoking (Figure 4.2, Panel A; Table 4.2). Targeting smoking cessation led to a -1.8% (95% CI: -3.6, 0; Risk Ratio [RR]: 0.83; 95% CI: 0.64, 1.0) reduction in the 8-year risk of mortality while targeting both alcohol and smoking risk behaviors led to a -2.0% (95% CI: -3.9, -0.2) reduction (RR: 0.81; 95% CI: 0.62, 0.99).

Results were attenuated with more realistic reductions in alcohol and smoking based on the efficacy of existing interventions (Figure 4.2, Panel B; Table 4.3). The 8-year risk of mortality was 10.1% (95% CI: 6.1, 14.2) with alcohol reduction, 9.5% (95% CI: 5.8, 13.3) with reduction in the smoking prevalence, and 9.3% (95% CI: 5.2, 13.0) with both alcohol and smoking reduction interventions. The combined alcohol and smoking reduction intervention reduced the risk of mortality by -1.1% (95% CI: -2.1, -0.1) (RR: 0.89; 95% CI: 0.81, 0.98).

E. Discussion

Proposed interventions on non-AIDS risk factors notably reduced the estimated risk of all-cause mortality among HIV-positive women enrolled in the WIHS. All interventions that eliminated or reduced alcohol consumption and/or smoking improved survival compared to intervening only on prompt initiation of ART during the modern treatment era, though confidence intervals for alcohol interventions did not exclude the null. Strategies that eliminated rather than reduced alcohol and smoking achieved the greatest decreases in the 8-year risk of all-cause mortality. Nearly eliminating alcohol and smoking reduced the 8-year risk of mortality by 20%.

While modern ART has transformed the prognosis of HIV to a manageable chronic condition, non-AIDS risk factors remain significant contributors to early mortality. Yet, it has been unclear what long-term improvements in survival may be achieved by interventions that combine initiation of ART with interventions on non-AIDS risk factors. Two recent studies used observational data from the WIHS to evaluate the reduction in the mortality risk that may be achieved by combining ART with depression treatment[74] and combining ART with treatment of hepatitis C infection.[65] Using the parametric g-

computation algorithm to estimate the risks of mortality under modeled interventions, the authors demonstrated improved survival in both settings.

Similar to depression and hepatitis C infection, alcohol consumption and smoking are common non-AIDS risk factors among people with HIV. The estimated prevalence of smoking and heavy alcohol use in those with HIV is over twice as high as in the general population.[23,24] Furthermore, these risk factors have especially pronounced adverse effects among those with HIV,[42,47] decrease the likelihood of viral suppression,[136,137] and negatively impact adherence to HIV treatment.[137,138] However, survival under interventions combining initiation of ART in the modern treatment era with interventions on alcohol and/or smoking have not been assessed or compared to determine which strategies should be prioritized.

Our finding that interventions on smoking in combination with modern ART achieved greater reductions in the 8-year risk of mortality than interventions on alcohol in combination with modern ART is likely influenced by the large proportion of smokers compared to drinkers in the WIHS. At baseline, there were over 2.5 times as many women reporting smoking than alcohol consumption of over 3 drinks per week. In fact, smoking was more common than consuming any alcohol or over 1 drink per week (59% versus 51% or 34%). Thus, interventions that reduced or eliminated smoking in the WIHS affected a greater proportion of the study sample than those that nearly eliminated or reduced drinking. In other HIV cohorts with varying distributions of these risk factors, the impacts of interventions eliminating or reducing smoking versus alcohol may differ. However, in the US HIV clinical setting, smoking is more common than at-risk alcohol use among HIV patients.[139]

While interventions on smoking as opposed to alcohol consumption in the WIHS achieved the greatest reductions in the mortality risk by 8 years, this finding was not consistent across the entire follow-up period. For interventions reducing the prevalence of alcohol and/or smoking, risks of all-cause mortality under each portfolio were similar for the first two years. After this time point, interventions on smoking (or both smoking and alcohol) combined with prompt initiation of modern ART progressively reduced the risk of mortality compared to intervening on modern ART alone. While the intervention on alcohol combined with ART increasingly improved survival over the remaining follow-up period, this

improvement was more modest. For interventions eliminating alcohol and smoking, the elimination of alcohol appeared slightly more effective during the first two years, after which the estimated risks of mortality progressively diverged to reveal greater improvements in survival with interventions on smoking (or both smoking and alcohol) than with interventions on alcohol use. It is possible that smoking cessation may take a longer time to affect the risk of mortality than changes in alcohol consumption. Thus, while eliminating or reducing alcohol may convey similar short-term benefits in the WIHS, smoking cessation is critical to achieve optimal long-term outcomes.

Though all intervention portfolios reduced the risk of all-cause mortality in the WIHS, near elimination of alcohol and/or smoking resulted in a two times greater reduction in the mortality risk as compared to interventions reflecting more realistic reductions in the prevalence of risk factors. Risks estimated under elimination strategies offer a promising benchmark for the improvements in survival that may be achieved by interventions with 100% efficacy. However, perfectly efficacious interventions against smoking and alcohol use do not currently exist.

Therefore, we estimated risks under Set B intervention portfolios to provide a more realistic view of what may be achieved with existing interventions.[46,49–51] The reduction in alcohol consumption and smoking prevalence that we modeled in the WIHS was based on efficacy estimates from actual interventions assessed in trials. The smoking cessation strategy we modeled in the WIHS was based on a computer expert system behavioral intervention delivered in physician waiting rooms that achieved only a 0.16 probability of smoking cessation by 6 months.[121] Yet, the low probability of sustained smoking cessation is similar to that estimated in other behavioral and nicotine replacement therapy studies[60,140] and highlights the need for more efficacious smoking cessation interventions among those with HIV.

We modeled the expected efficacy of alcohol reduction among those with hepatitis C virus based on a brief counseling intervention conducted in a hepatitis C clinic.[120] While the intervention achieved reductions in alcohol consumption among 62% of participants, it was not tested specifically among those with HIV. For those with HIV but not hepatitis C virus, most trials have been conducted among those with an indication of alcohol dependency or hazardous levels of consumption well beyond the national

recommended safe limits.[141–143] However, with potential for physiologic injury and increased mortality at lower thresholds of alcohol consumption among those with HIV,[47] it is necessary to reassess safe limits for this population and evaluate interventions tailored to reduce lower levels of drinking.

We made several assumptions in our analyses, including no measurement error of exposures, covariates, or outcome;[107] counterfactual consistency;[108,144] no unmeasured confounding;[135] positivity;[111] and correct specification of parametric models for g-computation. Information on alcohol and smoking were self-reported in the WIHS, and as such, may be underreported. In practice though, the provision of alcohol and smoking cessation interventions will likely depend on self-report of these risk behaviors. Unmeasured confounding is always a threat to observational studies; however, the rich data collected by the WIHS allowed us to address key baseline and time-varying confounders. In our models, we improved positivity (i.e., nonzero probability of exposures within all strata of covariates) by careful selection of covariates, functional form, and interaction terms to avoid cells with sparse data. We assessed specification of parametric models by comparisons of the modeled data distribution to that observed in the WIHS cohort (Figure 4.3). For alcohol and smoking interventions, we assumed no relevant side effects on risk of mortality through pathways other than those we modeled.[52] Further, due to the nature of the WIHS as an interval cohort with measures of ART, alcohol, and smoking documented at 6-month follow-up visits among those who remain alive, we assumed that any change in exposure necessitated at least a 6-month period to affect the risk of mortality.

Finally, there are several noteworthy strengths of our study. The WIHS is the largest interval cohort of HIV-positive women in the United States. Many studies conducted among those with HIV have been restricted to men or cohorts that are over 75-80% men such as the United States CFAR Network of Integrated Clinical Systems (https://www.uab.edu/cnics/).[63] However, women experience a different distribution of non-AIDS risk factors than men, including a higher prevalence of multimorbidity.[5] Evaluating the impact of interventions on non-AIDS risk factors among this group is necessary to improve health outcomes and prevent disparities from growing.[50,51] We also used the WIHS to compare multiple sets of interventions on non-AIDS risk factors. Few studies have examined sets of interventions

or compared interventions targeting different risk factors in a given HIV-positive population. Given the clustering of comorbidities among those with HIV, comparisons of interventions are critical to prioritize strategies for improved healthcare delivery.

F. Conclusion

Closing the remaining gap in the health status and survival of those with HIV in the modern ART era requires comprehensive approaches to HIV care management. Interventions that target smoking and alcohol may achieve significant reductions in the mortality risk of women with HIV who have received combination ART. To fully recognize these improvements, more efficacious interventions may be necessary, particularly to achieve higher proportions of smoking cessation.

Table 4.1. Characteristics of 1,016 HIV-positive combination antiretroviral therapy (ART) naïve women in the Women's Interagency HIV Study (WIHS) at analysis baseline.

Characteristics	
No. of women	1,016
Age, y, median [IQR]	38 [32, 45]
CD4 count, cells/mm^3, median [IQR] [a]	452 [273, 655]
Viral load, $\log_{10}$ copies/mL, median [IQR] [b]	3.6 [2.8, 4.4]
AIDS-defining illness	323 (31.8)
Average alcohol consumption	
Abstain	496 (48.8)
1 drink per week	174 (17.1)
2-3 drinks per week	117 (11.5)
4-7 drinks per week	81 (8.0)
7+ drinks per week [c]	148 (14.6)
Black race	696 (68.5)
Depressive symptoms [d]	513 (50.5)
Education	
Less than high school	381 (37.5)
High school graduate	328 (32.3)
Some college	307 (30.2)
Hepatitis C virus [e]	333 (32.8)
Hypertension [f]	257 (25.5)
Injection drug use history	338 (33.3)
Smoking status	
Current	595 (58.6)
Former	155 (15.3)

Data are presented as No. (%) unless otherwise specified.

[a] Missing for 4 women.

[b] Missing for 16 women.

[c] Above national recommended safe limits.

[d] Score $\geq$16 on the Center for Epidemiologic Studies Depression Scale (CES-D).

[e] Antibody positive and RNA positive or missing RNA value.

[f] Any indication of hypertension defined as systolic blood pressure $\geq$140, diastolic blood pressure $\geq$90, self-report of hypertension, or use of anti-hypertensive medications.

Table 4.2. Risk of all-cause mortality under hypothetical intervention portfolios to eliminate non-AIDS risk factors among 1,016 HIV-positive, combination ART-naïve women enrolled in the Women's Interagency HIV Study, 1998 – 2017.

Intervention		5-year risk, %	8-year risk, %	Risk Ratio (8 years)	Risk Difference (8 years)
	Natural course ART initiation	13.6	22.5		
(1)	Prompt initiation of modern ART [a]	7.4 (4.7, 10.1)	10.4 (6.3, 14.5)	1.00	0.0
(2a)	Strategy 1 & alcohol elimination [b]	7.1 (4.3, 9.8)	9.9 (5.9, 14.0)	0.95 (0.88, 1.02)	-0.5 (-1.2, 0.3)
(3a)	Strategy 1 & 100% smoking cessation	6.3 (3.7, 8.9)	8.6 (5.1, 12.1)	0.83 (0.64, 1.00)	-1.8 (-3.6, 0.0)
(4a)	All of the above	6.3 (3.5, 8.8)	8.4 (4.3, 11.9)	0.81 (0.62, 0.99)	-2.0 (-3.9, -0.2)

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus.
[a] Prompt initiation of modern ART was defined as initiation of ART in the treatment era after which tenofovir was available and within 6 months of the analysis baseline.
[b] The alcohol elimination intervention was specified as eliminating alcohol consumption >1 drink per week among women without hepatitis C virus and complete abstinence among women with hepatitis C virus.

Table 4.3. Risk of all-cause mortality under hypothetical intervention portfolios to reduce the prevalence of non-AIDS risk factors among 1,016 HIV-positive, combination ART-naïve women enrolled in the Women's Interagency HIV Study, 1998 – 2017.

Intervention		5-year risk, %	8-year risk, %	Risk Ratio (8 years)	Risk Difference (8 years)
	Natural course ART initiation	13.6	22.5		
(1)	Prompt initiation of modern ART [a]	7.4 (4.7, 10.1)	10.4 (6.3, 14.5)	1.00	0.0
(2b)	Strategy 1 & alcohol reduction [b]	7.3 (4.6, 10.0)	10.1 (6.1, 14.2)	0.98 (0.92, 1.03)	-0.3 (-0.9, 0.3)
(3b)	Strategy 1 & smoking cessation [c]	6.9 (4.3, 9.5)	9.5 (5.8, 13.3)	0.91 (0.84, 0.99)	-0.9 (-1.7, -0.1)
(4b)	All of the above	6.9 (4.1, 9.4)	9.3 (5.2, 13.0)	0.89 (0.81, 0.98)	-1.1 (-2.1, -0.1)

Abbreviations: ART, antiretroviral therapy; HIV, human immunodeficiency virus.
[a] Prompt initiation of modern ART was defined as initiation of ART in the treatment era after which tenofovir was available and within 6 months of the analysis baseline.
[b] The alcohol reduction intervention was defined as reducing alcohol intake to 3 drinks per week among those without hepatitis C virus; among those with hepatitis C virus, the intervention was defined as a 0.36 probability of abstaining from alcohol, 0.26 probability of reducing drinking levels by 50%, and a 0.38 probability of no change in alcohol consumption based on the intervention described by Dieperink et al.[120]
[c] The smoking cessation intervention was defined as a 0.16 probability of quitting smoking by the next follow-up visit based on the intervention described by Hoffman et al.[121]

Figure 4.1. Risk of all-cause mortality under four antiretroviral therapy initiation strategies in the Women's Interagency HIV Study, 1998 – 2017.

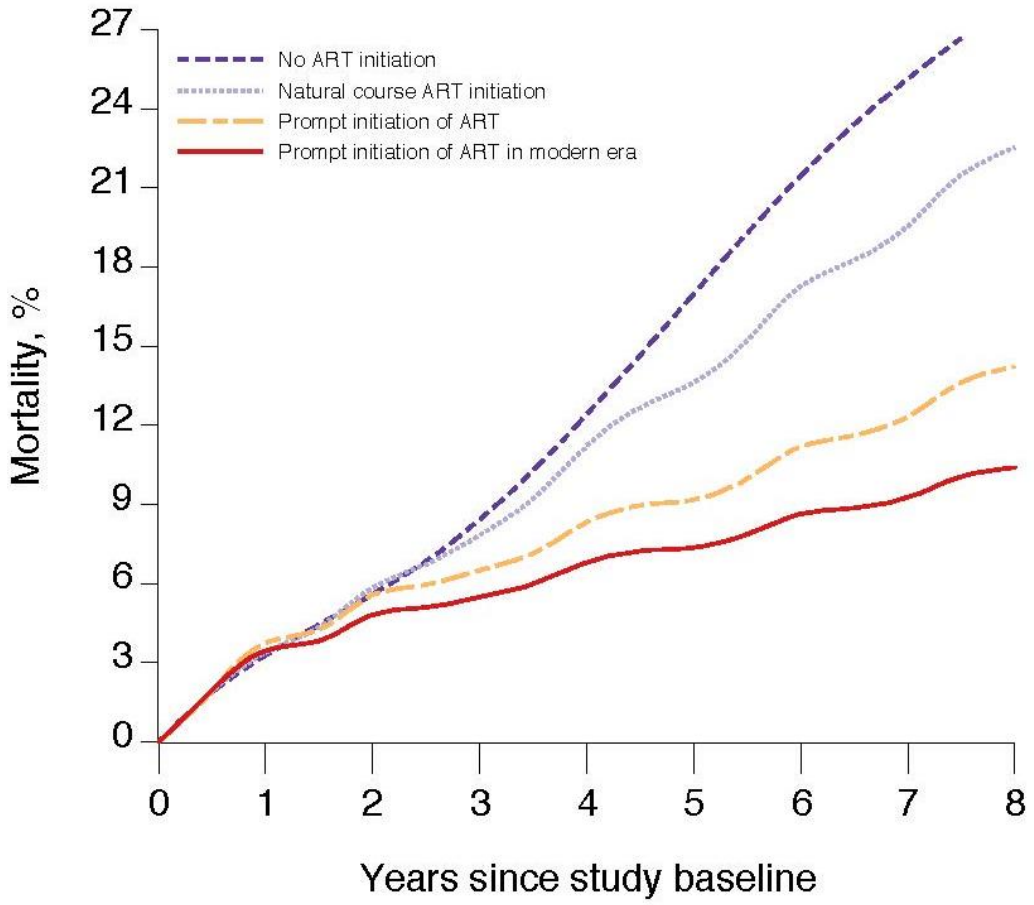
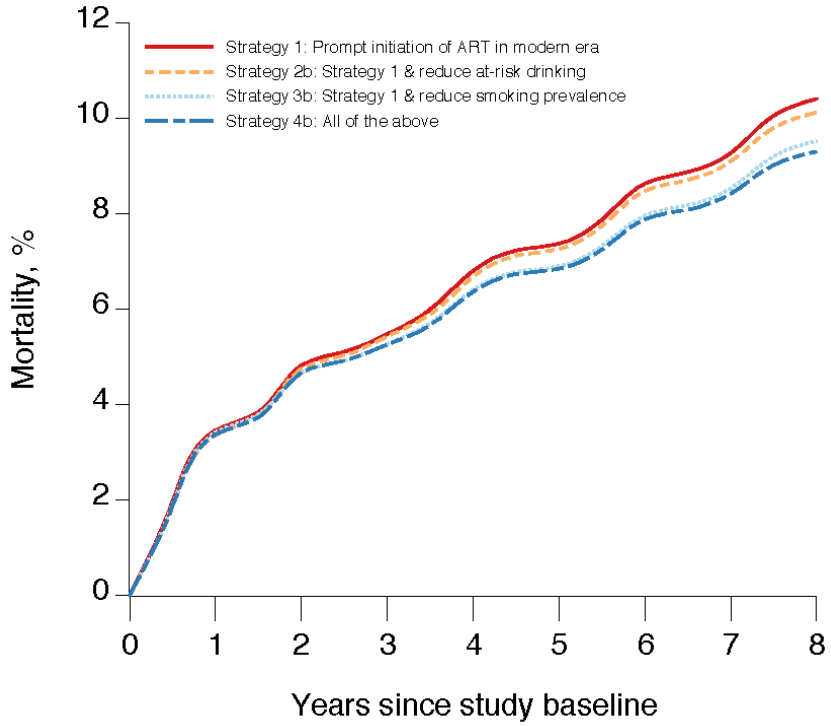
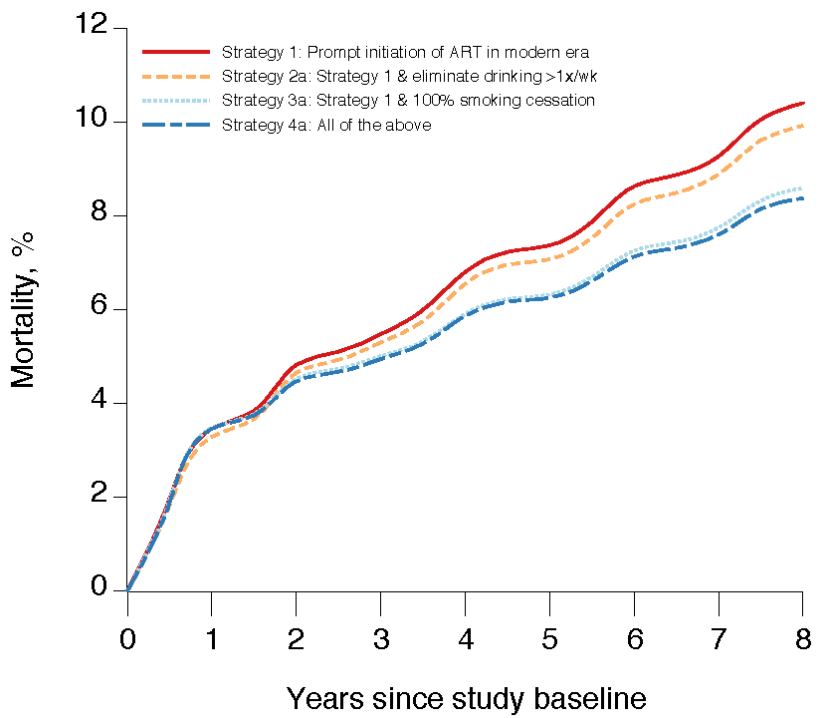


Figure 4.2. Risk of all-cause mortality under hypothetical interventions to eliminate (Panel A) or reduce (Panel B) the prevalence of non-AIDS risk factors in the Women's Interagency HIV Study, 1998 – 2017.

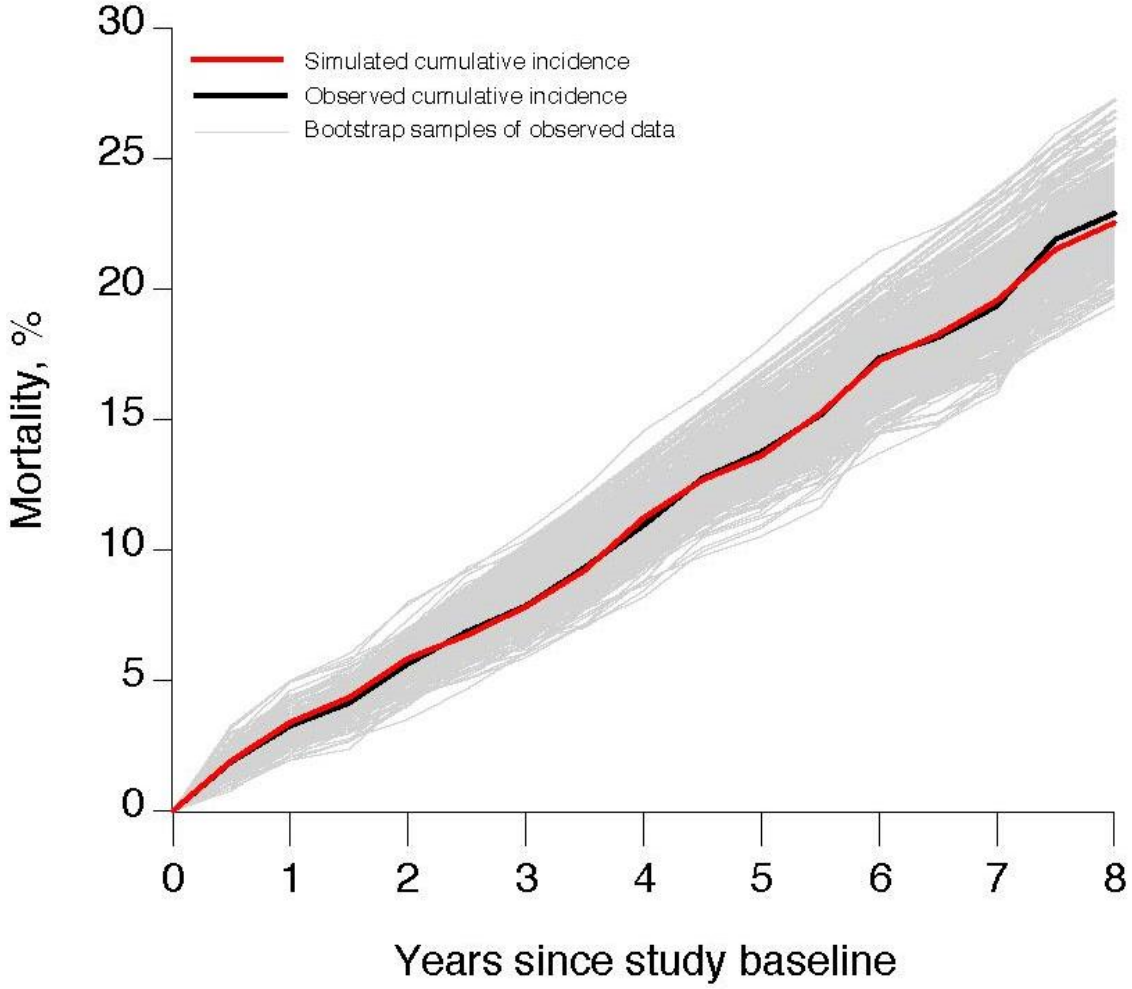


A)



B)

Figure 4.3. Observed versus g-computation simulated cumulative incidence of all-cause mortality under the natural course of antiretroviral therapy initiation among 1016 HIV-positive women enrolled in the Women's Interagency HIV Study between 1998 – 2017, 200 bootstrap samples of observed data.



CHAPTER 5: TWO-STAGE G-COMPUTATION: EVALUATING TREATMENT AND INTERVENTION IMPACTS IN OBSERVATIONAL COHORTS WHEN EXPOSURE INFORMATION IS PARTLY MISSING

A. Overview

A barrier to implementing the g-computation algorithm to evaluate average treatment effects and intervention impacts using observational data is that the approach assumes complete information on exposures. We develop two-stage g-computation estimators that leverage partially observed information on the full study sample and complete exposure information on a subset. In a hypothetical cohort of 1623 HIV-positive women with 30% complete opioid prescription information, we illustrate a two-stage extrapolation g-computation estimator for the average treatment effect of shortening versus lengthening the duration of all opioid prescriptions; we further illustrate two-stage inverse probability weighted and exposure imputation g-computation estimators for the average intervention effect of shortening the duration of all prescriptions versus the status quo. Two-stage extrapolation g-computation approximated the true -18.6% risk difference for the average treatment effect while two-stage inverse probability weighted and exposure imputation g-computation estimators approximated the true -10.4% risk difference for the average intervention effect. For both parameters, g-computation fit to the subset of complete cases was biased. In 10,000 Monte Carlo simulations, two-stage approaches considerably reduced bias and mean squared error and improved 95% confidence limit coverage. While missing data threaten validity and precision, proposed two-stage g-computation designs can be used to make progress in the face of these challenges.

B. Introduction

The generalized computation formula algorithm[67] (g-formula) is increasingly applied to observational data to answer pressing public health questions, especially when conducting a randomized

controlled trial may be costly, inefficient, or infeasible. Under a set of conditions,[107,109–111,145] the g-formula re-expresses the data distribution under the observed (or factual) exposure plan as the distribution that would have been observed under an alternative (or counterfactual) exposure plan. Recent papers have illustrated how the g-formula can be estimated parametrically using available statistical software,[68,69,71,73,98] with applications highlighting the versatility of the g-formula in estimating both: 1) the "treatment/exposure effect" (i.e., a contrast of the average outcome distribution when the full population is exposed versus unexposed) and 2) "intervention effects" (i.e., contrasts of the population average outcome distribution under a realistic change in the exposure distribution versus no change).[46,49,88,103] Intervention effect estimates are particularly important to inform clinical guidelines and policy decisions.[50,51]

Applications of the parametric g-formula assume complete information on exposures, but missing data are extensive in observational studies.[77,81] A key barrier to implementing this powerful analytic tool in many research settings is the lack of high quality data on exposures or interventions in the entire study sample. Exposures may not be observed in the full sample when information is costly or difficult to measure on everyone, data linkages are only performed for a subset, or information collected has changed across time or study sites.

As Lash and Schisterman noted in their call for "New Designs for New Epidemiology," answering important scientific questions in settings with constrained resources and fragmented real-world data requires novel study designs.[146] Two-stage sampling designs formalized by Neyman[82] and developed in traditional regression analyses[80,83–87,96] have improved efficiency and validity of estimators when information on an important variable is incomplete among the total sample; however, analogous extensions for comparing counterfactual outcome distributions with the parametric g-computation algorithm when exposure is partially missing have yet to be developed.

Here, we extend two-stage analytic design and missing data methods to develop two-stage g-computation estimators of the average treatment and intervention effects that leverage exposure information on a subset of participants along with covariate information observed for the full study

sample. We propose and implement estimators in an example motivated by a hypothetical HIV-positive cohort in which information on exposure to the candidate intervention is only available for a subset. We assess the operating characteristics of the proposed estimators with simulation and evaluate estimators under varied sample conditions.

C. Methods

C1. Motivating Example Data

Our objectives are to estimate the average treatment and intervention effects[46,49,88,103] of opioid prescription duration among HIV-positive women on the 12-month risk of emergency room/hospital visit or death from any cause. Specifically, our goal is to estimate the average treatment effect of shortening versus lengthening the duration of initial opioid prescriptions among all women and the average intervention effect of shortening the duration of initial prescriptions versus status quo prescribing practices. To illustrate two-stage estimation of parameters, we base our example on a hypothetical cohort simulated to represent HIV-positive women enrolled in the Women's Interagency HIV Study (WIHS). [89–91] The WIHS is a United States-based prospective interval cohort study of HIV-positive and HIV-negative women enrolled beginning in 1994 and followed at semiannual study visits for behavioral and clinical data. We created our dataset such that the full sample size was equal to the number of HIV-positive women enrolled in Waves 2 – 4 of the WIHS (i.e., $N = 1623$) and the incidence of emergency room/hospital visit or death over 12 months was similar to that observed in a sample of outcome data from the most recent wave of enrollees and which has been reported in previous HIV studies[130] (i.e., approximately 30%).

While information on opioid use has been collected on WIHS participants, this information has primarily been based on self-reported use of illicit or recreational drugs. Our objective is to estimate the effect of a clinical intervention in which guidelines dictate that physicians shorten the period over which initial opioid prescriptions are provided. Linkages to electronic medical record databases could provide detailed dosage and duration information for individuals receiving prescriptions; however, because data linkages are expensive and often cannot be performed for the entire study sample, we consider a scenario

in which it is possible to link prescription information for a nonrandom subset of 30% of the simulated cohort, particularly members receiving HIV care at 10 clinics proximal to the WIHS sites.

A causal diagram[122] for our research scenario shown in Figure 5.1 guided our generation of data for analyses. Receipt of an initial long duration opioid prescription at baseline (i.e., A) increases the risk of emergency room/hospital visit or death over 12-months (i.e., Y). History of substance use (i.e., W) represents one potential confounder increasing the probability of receiving an initial long duration prescription and the risk of emergency room/hospital visit or death. Not receiving regular HIV care (i.e., Z), while not a confounder and therefore often omitted from the causal diagram, decreases the probability of having observed exposure information (i.e., $R = 1$) from linked electronic medical records and also increases the risk of hospitalization or death (i.e., Y). Complete details on the data generation process are available in Appendix B of the Technical Supplement.

C2. G-computation

Let A denote a binary time-fixed exposure, $\boldsymbol{W}$ denote a vector of confounders, Y denote an outcome, uppercase letters denote random variables, and lowercase letters denote possible values of variables or constants. To simplify notation, we assume $\boldsymbol{W}$ to be a finite dimension vector of categorical covariates; in practice, the vector may include continuous covariates in which case the equations would be rewritten as integrals. The expected value of the potential outcome Y^a under a time-fixed exposure plan a can be written as,

$$(2.1) \quad E[Y^a] = \sum_{\boldsymbol{w}} E(Y|A = a, \boldsymbol{W} = \boldsymbol{w})P(\boldsymbol{W} = \boldsymbol{w})$$

where we define $E[Y^{a=1}]$ as the expected outcome distribution under a plan in which the full study sample receives exposure and $E[Y^{a=0}]$ as the expected outcome distribution under a plan in which exposure is removed from the full study sample. The expected value of the outcome under the natural course (i.e., observed exposure plan) can be written as,

$$(2.2) \quad E[Y] = \sum_{\boldsymbol{w},a} E(Y|A = a, \boldsymbol{W} = \boldsymbol{w})P(A = a|\boldsymbol{W} = \boldsymbol{w})P(\boldsymbol{W} = \boldsymbol{w}).$$

The process of estimating the above expressions (referred to as the g-formula) by g-computation has been described in detail.[68,69,71,98] Briefly, this is performed by first estimating the conditional probabilities or densities of the outcome given exposure and confounders from models fit to the full study sample. Next, a large Monte Carlo sample is drawn from the original study population. The exposure values for each individual in the Monte Carlo sample are set according to the exposure plan. Then, outcomes are simulated in the Monte Carlo sample using the set exposure value a, realized covariate values $\boldsymbol{w}$, and conditional probabilities that were estimated in the first step. For each individual, the exposure value set by plan a may differ from the observed exposure value. Finally, outcomes are summed across the distribution of $\boldsymbol{w}$. In special cases, the Monte Carlo simulation step may be avoided.[70] It is standard practice to estimate the outcome distribution under the natural course with g-computation regardless of whether it is a parameter of interest; comparison of the observed natural course to the g-computation simulated natural course is used as a check on model specification prior to using the models to simulate outcome distributions under altered exposure plans.

When exposure plan information for some participants is missing, missing data prohibit the use of the above approach. Instead, under an assumption that patients with (i.e., $R = 1$) and without (i.e., $R = 0$) complete exposure information are exchangeable within levels of confounders W, investigators often use a complete case analysis. Under this assumption, the g-formula equations for the potential outcome under exposure plan a and under the natural course can be written as:

$$(2.3) \quad E[Y^a] = \sum_{\boldsymbol{w}} E(Y|A = a,\ \boldsymbol{W} = \boldsymbol{w}, R = 1)P(\boldsymbol{W} = \boldsymbol{w}, R = 1)$$

$$(2.4) \quad E[Y] = \sum_{\boldsymbol{w},a} E(Y|A = a,\ \boldsymbol{W} = \boldsymbol{w}, R = 1)P(A = a|\boldsymbol{W} = \boldsymbol{w}, R = 1)P(\boldsymbol{W} = \boldsymbol{w}, R = 1).$$

The estimation of conditional outcome probabilities or densities is restricted to the subsample with observed exposure (i.e., complete information), and the outcome distribution is summed over the distribution of $\boldsymbol{w}$ among those with complete information. However, $E[Y^a|W = w]$ will only be

approximated by $E[Y^a|W = w, R = 1]$ if exposure information is missing completely at random or missing at random given the vector of confounders $\boldsymbol{W}$.

C3. Two-stage G-computation Estimator of Average Treatment Effect $\boldsymbol{E[Y^{a=0} - Y^{a=1}]}$

We propose a g-computation extrapolation estimator of the average treatment effect which combines partial information on the full study sample with complete information on the subset of participants with observed exposure. We rewrite the g-formula to 1) include $\boldsymbol{Z}$ in our outcome model, where $\boldsymbol{Z}$ is a vector of covariates associated with R and Y that allows us to meet a weaker assumption that presence of exposure information is independent of Y given $\boldsymbol{W}$ and $\boldsymbol{Z}$ (i.e., $E[Y^a|a, \boldsymbol{w}, \boldsymbol{z}] = E[Y^a|a, \boldsymbol{w}, \boldsymbol{z}, R = 1]$), and 2) sum outcomes over the joint distribution of $\boldsymbol{W}$and $\boldsymbol{Z}$ instead of only $\boldsymbol{W}$ in the subset with complete exposure information. The g-computation extrapolation estimator is written as:

$$(3.1) \quad \hat{E}_{ext}[Y^0 - Y^1] =$$

$$\sum_{\boldsymbol{w},\boldsymbol{z}} \hat{P}(Y = 1|A = 0, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)\hat{P}(\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z})$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z}} \hat{P}(Y = 1|A = 1, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)\hat{P}(\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}) .$$

Stage 1 consists of estimating $\hat{P}(Y = 1|A = a, \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, R = 1)$ from a logistic model fit to complete cases; stage 2 consists of (a) simulating outcomes under plans $a = 0$ and $a = 1$ among Monte Carlo resamples taken from the *full* population (i.e., not only sampling complete cases) and then (b) summing over the full population's joint distribution of $\boldsymbol{w}$ and $\boldsymbol{z}$. We obtain the average treatment effect by taking the difference $\hat{E}[Y^0 - Y^1]$.

C4. Two-stage G-computation Estimators of Average Intervention Effect $\boldsymbol{E[Y^{a=0} - Y]}$

The risk under the natural course is estimated by additionally summing over the full population's distribution of observed exposure. Because the intervention effect contrasts the outcome distribution under an altered exposure distribution to the outcome distribution under the observed exposure distribution, an additional step is needed when exposure status is missing for a subset of the sample and the outcome distribution differs among the subset with observed information versus the full sample.

Similarly, this is needed for interventions that depend on the natural value of exposure[104] such as for the effect estimated by Lesko et al.[64] in which intervening on depression depended on the natural value (i.e., observed level) of depressive symptoms.

We propose two-stage inverse probability-weighted (IPW) and imputation g-computation estimators of the average intervention effect. As with the extrapolation g-computation estimator, IPW and imputation g-computation include $\boldsymbol{Z}$ in the outcome model as well as the additional model specified for missing data.

The IPW g-computation estimator is written as:

(4.1) $\hat{E}_{IPW}[Y^0 - Y] =$

$$\sum_{\boldsymbol{w},\boldsymbol{z}}\left[\frac{\hat{P}(Y=1|A=0,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z}|R=1)}{\hat{P}(R=1|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})}\right]$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z},a}\left[\frac{\hat{P}(Y=1|A=a,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(A=a|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z}|R=1)}{\hat{P}(R=1|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})}\right].$$

Stage 1 consists of estimating the probability of being a complete case (i.e., $\hat{P}(R=1\,|\,\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$)[131] in the *full* population in addition to estimating $\hat{P}(Y=1|A=a,\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},R=1)$ in the complete cases as in the extrapolation approach. In stage 2, we (a) simulate outcomes setting $a=0$ and $A=a$ under the observed exposure value among Monte Carlo resamples taken from the *complete cases*, (b) weight simulated outcomes by the inverse probability of being a complete case, (c) and sum over the joint distribution of $\boldsymbol{W}$ and $\boldsymbol{Z}$ in the *complete cases.*

The imputation g-computation estimator is written as:

(4.2) $\hat{E}_{impute}[Y^0 - Y] =$

$$\sum_{\boldsymbol{w},\boldsymbol{z}}\hat{P}(Y=1|A=0,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$$

$$-\sum_{\boldsymbol{w},\boldsymbol{z},a}\hat{P}(Y=1|A=a,\ \boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z},\ R=1)\hat{P}(A^*=a|\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})\hat{P}(\boldsymbol{W}=\boldsymbol{w},\boldsymbol{Z}=\boldsymbol{z})$$

where a denotes the realized value of observed exposure or imputed exposure (where missing). A separate model is specified in stage 1 to estimate the conditional probability of exposure among those with complete information, (i.e., $\hat{P}(A = 1|\boldsymbol{W} = \boldsymbol{w}, \boldsymbol{Z} = \boldsymbol{z}, Y = y,\ R = 1)$); in stage 2, (a) values of exposure are imputed for each woman with missing exposure information in the Monte Carlo resamples taken from the *full* population using the conditional probabilities estimated in stage 1, and (b) outcomes are simulated under the *full* joint distribution of imputed or observed a^*, $\boldsymbol{w}$, and $\boldsymbol{z}$. Because no values for a are imputed when a is set to 0, the first part of the equation for Y^0 in equation 5.1 is identical to that of the extrapolation approach in 3.1.

C5. Analyses of Example Cohort

We used proposed two-stage g-computation estimators as well as the naïve, complete case g-computation estimator to estimate parameters of interest among our simulated cohort of 1623 HIV-positive women. We estimated 12-month absolute risks of emergency room/hospital visit or death if all women were to receive short duration opioid prescriptions for ≤7 days (i.e., $\hat{E}[Y^{a=0}]$ if all women were to receive long duration opioid prescriptions for >7 days (i.e., $\hat{E}[Y^{a=1}]$), and if prescription practices were to remain unchanged (i.e., $\hat{E}[Y]$). We estimated risk differences contrasting the average treatment effect if all women received short duration versus long duration opioid prescriptions as well as the average intervention effect if all women received short duration opioid prescriptions versus the status quo.

Because two-stage IPW and imputation g-computation can also estimate the average treatment effect, we used them in addition to the extrapolation estimator to compare approaches. In all analyses, we used a Firth correction[132] to improve model convergence. We estimated the standard error based on the standard deviation of 200 nonparametric bootstrap resamples to obtain 95% confidence intervals.[126] To propagate the error in the estimation of IP weights and exposure imputation, weight estimation and imputation were performed within bootstrap resamples.

C6. Simulation Experiments

To assess estimator operating characteristics, we repeated analyses of the example cohort in 10,000 simulation experiments performed under the same conditions. Additionally, we varied the percentage of complete exposure data from 10-90%. Using the proposed g-computation approaches, we estimated absolute risks and risk differences as the average point estimate across the 10,000 experiments. The true risks and risk differences were determined using the potential outcomes $Y^{a=1}$, $Y^{a=0}$, and Y from the full cohort prior to imposing missing exposure information. Bias was defined as the difference between the estimated and true risk or risk difference. Root mean squared error was computed as the square root of the sum of the squared bias and variance. 95% confidence limit coverage was defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation experiments.

Furthermore, we conducted 10,000 simulation experiments on four cohorts generated under scenarios in which all but one of the conditions remained the same as in our motivating example. In the first scenario, to confirm that estimators are valid, we simulated the cohort under the null opioid treatment and intervention effects, meaning that the risk under shortened, lengthened, and natural course opioid durations were equal and the risk differences were 0. In the second scenario, we generated the cohort under a case in which opioid information was missing completely at random to confirm standard g-computation works in this restrictive setting. In the third scenario, to assess performance of two-stage versus complete case approaches in the presence of strong modification, we generated the cohort under a case in which there was stronger modification of the opioid intervention effect by Z than in the base case. Finally, in the fourth scenario, to minimize finite sample bias, we increased the cohort size from 1,623 to 7,500 women.

D. Results

D1. Motivating Example Cohort

Table 5.1 presents the observed and complete data from one draw of 1623 HIV-positive women from our simulation experiments, of whom 30% (n = 512) had complete information on opioid

prescription duration. In the full cohort and the sample with complete exposure data, about 46% of women had a prior history of substance use. The prevalence of substance use history was higher among those receiving long duration versus short duration opioid prescriptions (53% versus 34%). About 42% of the full cohort was not engaged in regular HIV care compared to only 14% of those with complete data, due to lack of engagement in regular HIV care decreasing the likelihood of successful prescription record linkages. The 12-month risk of emergency room/hospital visit or death was higher in the full cohort (31%) compared to the sample with complete data (21%).

Table 5.2 presents the 12-month risks and risk differences for emergency room/hospital visit or death estimated by each of the g-computation approaches. The true risk of emergency room visit or death was 38.9% if all women received long duration opioid prescriptions, 20.3% if all women received short duration opioid prescriptions, and 30.7% under the natural course (i.e., status quo) in which 60% of women received long duration prescriptions and 40% received short duration prescriptions. Estimates obtained with the naïve complete case g-computation approach were biased downward and did not include the true value in the 95% confidence limits with the exception of the absolute risk under short duration prescriptions. The IPW and imputation g-computation estimators addressed missing values of opioid prescription duration to recover the risk under the natural course.

D2. Simulation Experiments

We repeated simulation and analyses of a cohort of 1623 HIV-positive women in 10,000 experiments. Figure 5.2 displays the risk differences estimated across 10,000 simulation experiments with 30% complete opioid prescription data. The true risk difference contrasting the provision of short duration versus long duration opioid prescriptions to all women (i.e., the average treatment effect) was -18.1%. All two-stage g-computation approaches except the complete case g-computation estimator yielded valid estimates. The IPW and imputation estimators, which can additionally be used to estimate the average intervention effect, performed well in estimating the true -11.1% reduction in the risk of emergency room/hospital visit or death.

The additional models specified for two-stage estimators of the average treatment effect resulted in larger standard errors than the complete case approach (Extrapolation SE: 5.1; IPW SE: 5.2; Imputation SE: 5.1; Complete Case SE: 3.7; Table 5.3a) as did the IPW and imputation g-computation estimators of the average intervention effect (IPW SE: 3.4; Imputation SE: 3.3; Complete Case SE: 2.3; Table 5.3b); however, root mean squared error was lower with two-stage estimation of the average treatment and intervention effects due to the large reduction in bias achieved by these estimators; confidence limit coverage was 94%. Two-stage approaches yielded estimates that were less biased than those from the complete case approach across all levels of complete data, and standard error decreased with increasing percentage of complete data (Figures 5.3 and 5.4).

Two-stage approaches performed well under the four scenarios incorporating null treatment and intervention effects, data missing completely at random, strong effect measure modification by Z, and large sample size (Tables 5.4a and 5.4b). When data were missing completely at random, the complete case g-computation estimator yielded valid estimates of the average treatment and intervention risk differences as expected, though there was a slight improvement in standard error with two-stage approaches as a result of the additional covariate information leveraged from the full sample. In all other scenarios, the complete case g-computation estimator failed to provide valid estimates.

E. Discussion

In our hypothetical HIV-positive cohort and varied simulation scenarios, two-stage g-computation estimators successfully leveraged completely observed exposure information on a subset of the cohort along with partially observed information on the full sample to provide valid population-level estimates of parameters of interest. Analogous to existing two-stage sampling designs for traditional regression estimators of conditional parameters,[86,87,96] two-stage g-computation estimators of risks and risk differences reduced bias compared to the traditional g-computation estimator fit to complete cases.

In all two-stage g-computation approaches, we included the covariate Z in the outcome model. Although Z was not a confounder, it was a predictor of complete information and the outcome. Consequently, conditioning on Z in the outcome model was necessary for d-separation of completeness

and the outcome,[147] as well as the models for IPW or exposure imputation to take into account the missing information.[148] When Z was left out of the outcome model, all two-stage approaches failed to provide unbiased effect estimates (Appendix B of the Supplemental Text; aFigure 1). Planning of studies should therefore include data collection on not only potential confounders but also additional covariates that are likely to predict both the outcome and having complete exposure.

It should be noted that we correctly specified IPW and imputation models for missing data in our simulation experiments. In determining which approach to use to estimate the average intervention effect, it will be important to consider one's ability to correctly model either the mechanism of complete information (for IPW) or the missing exposure values (for imputation). In practice, investigators might have better insight into one of these models. In validation studies relying on protocols for sampling a subset of the full population for improved exposure measurement, knowledge of the sampling mechanism can be leveraged in IPW. Alternatively, in circumstances where treatment is measured haphazardly, it might be easier to specify the imputation model.

Although two-stage approaches reduced bias across all levels of complete exposure data compared to the complete case estimator of average treatment and intervention effects, point estimates for the risk differences were not unbiased at lower percentages of complete data. This is likely due to finite sample bias imposed by the sample conditions of our simulation scenario. Bias decreased with decreasing percentage of missing data, and this pattern was less pronounced with a larger cohort size of $N = 7500$ (Appendix B of the Supplemental Text; aFigure 2). Nonetheless, even at low percentages of complete data in the original cohort of $N = 1{,}623$, two-stage approaches dramatically reduced bias compared to the complete case approach.

Because we generated our original cohort of N = 1623 with a 40% (n = 649) prevalence of short duration prescriptions and a 19% (n = 123) marginal risk of the outcome in this group, the number of observed events became small at low percentages of complete data (e.g., 10% complete data: about 12 events; 20% complete data: about 25 events; 30% complete data: about 37 events). Thus, in some of these scenarios, we had fewer than 10 events per variable in the model. Though we used a Firth correction[132] to

improve model convergence, consideration of the number of events within strata of exposure and covariates is important in analyses and might be used to motivate sampling designs that enrich for the outcome and covariates needed to ensure positivity. The covariates used to define these sampling designs would then become a part of the vector $\boldsymbol{Z}$ used to address partially observed data in two-stage g-computation estimation of parameters of interest. Further exploration of finite-sample properties of complete case and two-stage g-computation estimators would guide investigators in best practices for study planning and design.

We estimated a special case of an intervention effect in which it was possible to shorten the duration of initial opioid prescriptions among all HIV-positive women.[51] An example of an intervention that achieves this might be a change in electronic medical systems that prohibits physicians from submitting an opioid prescription for a patient that exceeds 7 days. However, in many clinical and public health policy settings, it might not be possible or desirable to completely eliminate an exposure through intervention. For example, in our study, we might also consider the same intervention to prohibit prescriptions exceeding 7 days but modified with an option for the physician to override this if patients are receiving the prescription for recovery from a specific type of surgical procedure. In this type of targeted dynamic intervention,[51] we could similarly use two-stage IPW or imputation g-computation estimators to recover the risk of the outcome under interventions that eliminate exposure among specific subgroups.

Finally, the proposed two-stage g-computation approaches for missing information are also methods to estimate the g-formula for generalizability.[149–152] The causal structure considered in this paper is identical to that which is considered in generalizability studies; the subset of those with complete information in our example (i.e., $R = 1$) is akin to those selected for the study (i.e., $S = 1$), while those with partially observed information in the full sample are akin to those with partially observed information in the target population. Therefore, two-stage designs could be applied to generalize estimated effects from a study sample to a target population of interest if information on the context (i.e., $\boldsymbol{W}$ and $\boldsymbol{Z}$) is available.

F. Conclusion

Advancements in methods and available statistical software programs as well as increasing access to fragmented, real-world data provide new opportunities to answer pressing public health questions among vulnerable populations. While missing data threaten validity and precision of estimates, two-stage study sampling designs have been employed in traditional regression analyses to overcome these challenges. Our study demonstrates that such designs can be extended to parametric g-computation estimation to quantify impacts of potential interventions and make progress in the face of these complexities.

Figure 5.1. Causal diagram used to generate a hypothetical cohort of 1,623 HIV-positive women where A represents initial opioid prescription duration, Y represents emergency room visit or death from any cause, W represents history of substance use, Z represents not being established in regular HIV care, and R is an indicator of having observed opioid prescription information with $R = 1$ indicating observed status and, therefore, selection into the sample of complete cases.

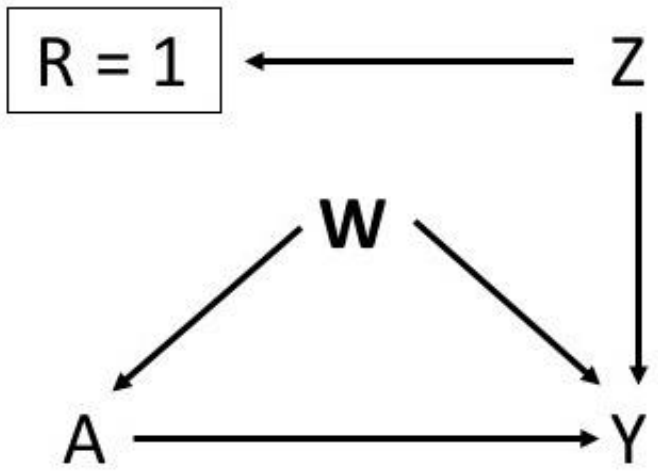


Figure 5.2. Risk difference estimated by g-computation approach contrasting provision of short duration versus long duration opioid prescriptions (Panel A) and provision of short duration prescriptions versus the status quo prescribing practice (Panel B) among 1,623 HIV-positive women with 30% complete opioid information, 10,000 simulation experiments.

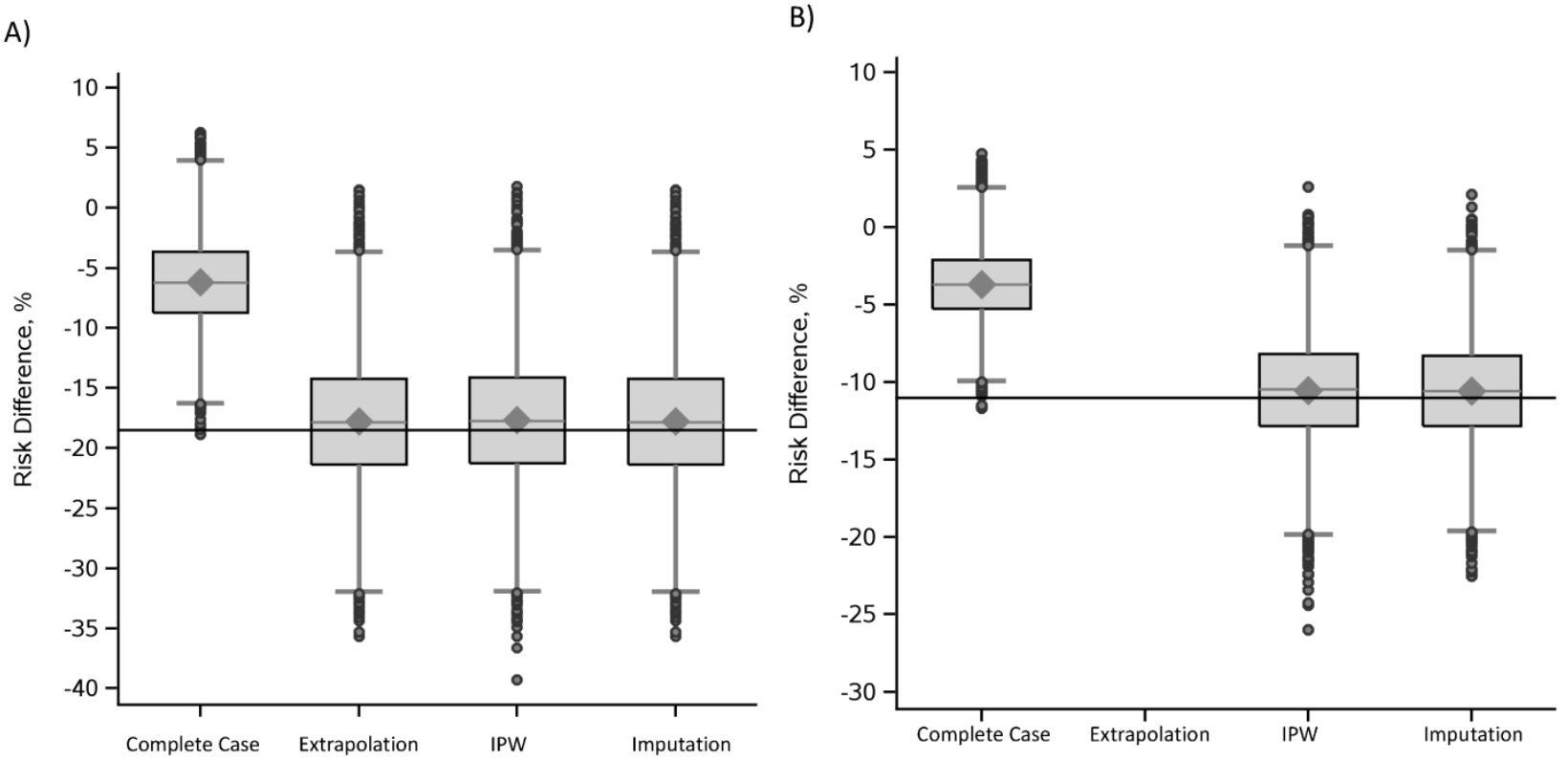


Figure 5.3. Bias (Panel A) and standard error (Panel B) for the average treatment effect of shortening versus lengthening opioid prescription duration among 1,623 HIV-positive women estimated by g-computation approach to address missing opioid data, 10,000 simulation experiments with complete opioid data varying from 10-90%.

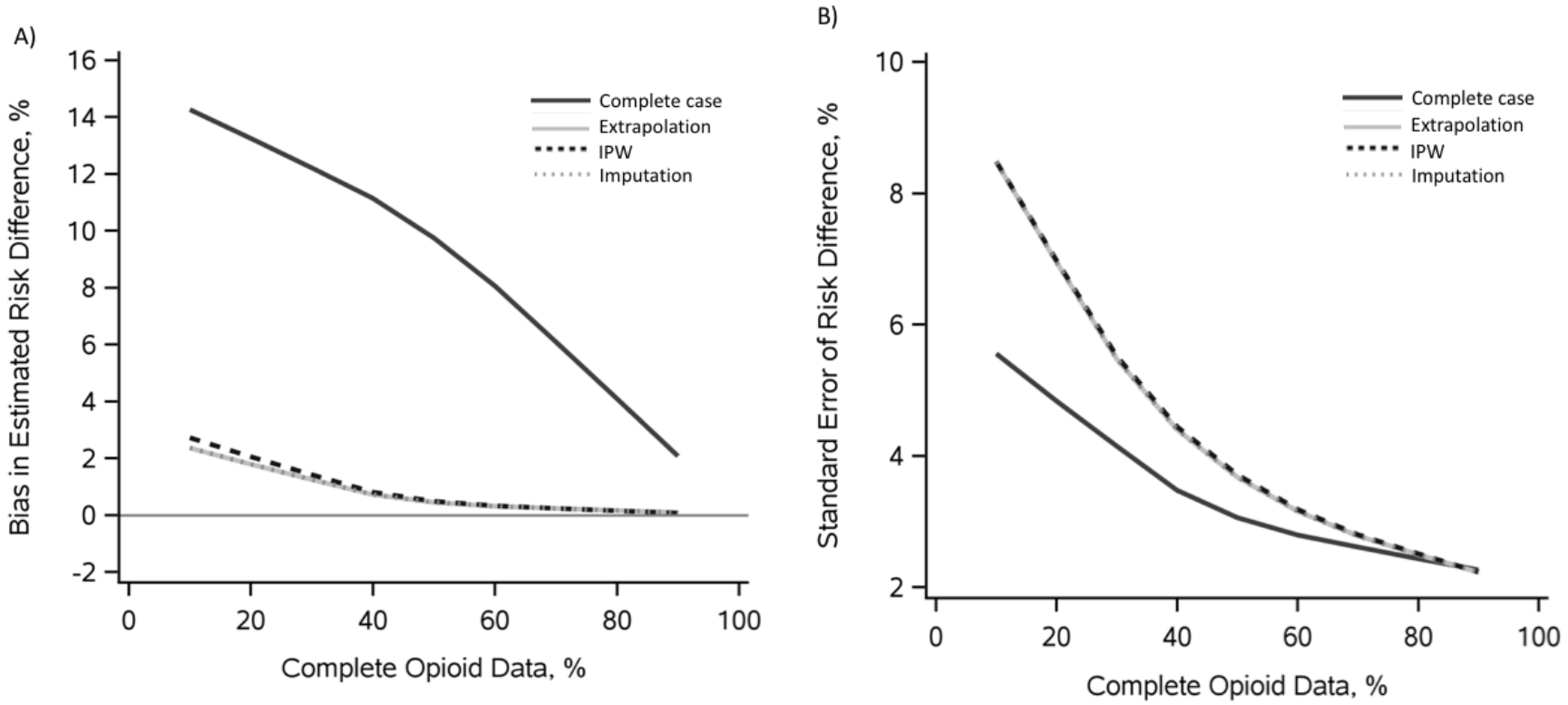


Figure 5.4. Bias (Panel A) and standard error (Panel B) for the average intervention effect of shortening opioid prescription duration versus the status quo prescribing practice among 1,623 HIV-positive women estimated by g-computation approach to address missing opioid data, 10,000 simulation experiments with complete opioid data varying from 10-90%.

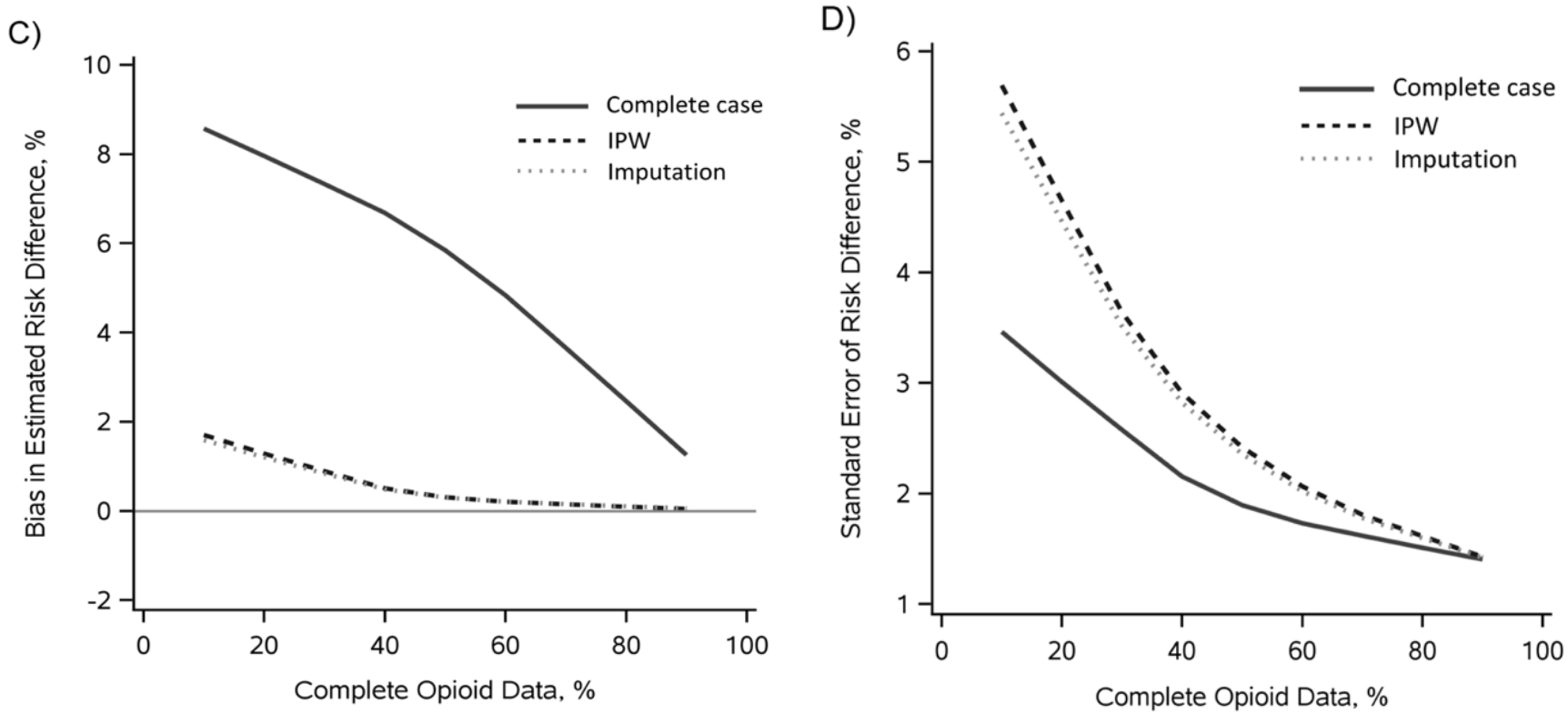


Table 5.1. Observed and complete data for a simulated cohort of 1,623 HIV-positive women. [a]

Characteristic	Full Cohort		Sample with Complete Exposure Data					
	Overall (N = 1,623)		Overall (n = 512)		Opioid RX 7+ days (n = 337)		Opioid RX ≤7 days (n = 175)	
	No.	%	No.	%	No.	%	No.	%
Substance use history	747	46.0	238	46.5	178	52.8	60	34.3
Not receiving regular HIV care	674	41.5	70	13.7	41	12.3	29	16.6
Emergency room visit or death	498	30.7	109	21.3	83	24.6	26	14.9

[a] One draw from our simulation scenario when complete opioid information is available for 30% of the cohort.

Table 5.2. Estimates of the twelve-month risks (95% confidence intervals)[a] of emergency room visit or death in a cohort of 1,623 HIV-positive women, by g-computation approach to address missing opioid information. [b]

Prescription duration	Truth	Complete Case	Extrapolation	IPW	Imputation
Absolute risks, %					
>7 days	38.9	24.6 (20.4, 28.9)	39.2 (32.6, 45.8)	39.1 (32.1, 46.1)	39.2 (32.6, 45.8)
≤7 days	20.3	17.2 (11.3, 23.1)	20.7 (12.9, 28.5)	20.2 (12.6, 27.9)	20.7 (12.9, 28.5)
Status quo	30.7	21.6 (18.1, 25.0)	--	30.4 (24.6, 36.3)	30.7 (25.9, 36.1)
Risk differences, %					
≤7 vs. >7 days	-18.6	-7.4 (-14.6, -0.2)	-18.5 (-28.4, -8.6)	-18.9 (-28.6, -9.2)	-18.5 (-28.4, -8.6)
≤7 days vs. status quo	-10.4	-4.4 (-9.4, 0.7)	--	-10.2 (-16.6, -3.8)	-10.3 (-16.8, -3.9)

Abbreviations: IPW, inverse-probability weighting.

[a] Standard error calculated as the standard deviation of the point estimate across 200 nonparametric bootstrap samples.

[b] Based on one draw from our simulation scenario with the study characteristics provided in Table 1 and where complete opioid information is available for 30% of the cohort.

Table 5.3a. Performance of g-computation approaches addressing missing data to estimate average treatment effect of shortening versus lengthening opioid prescription duration among a cohort of 1,623 HIV-positive women, 10,000 simulation experiments. [a]

	Complete Case		Extrapolation		IPW		Imputation	
	Estimate	Monte Carlo error	Estimate	Monte Carlo error	Estimate	Monte Carlo error	Estimate	Monte Carlo error
Bias [b]	12.4	(0.037)	0.8	(0.052)	0.9	(0.053)	0.8	(0.052)
Standard deviation of bias [c]	3.7		5.2		5.3		5.2	
Average standard error [d]	3.7	(0.003)	5.1	(0.005)	5.2	(0.005)	5.1	(0.005)
Mean squared error [e]	1.8	(0.010)	0.5	(0.004)	0.6	(0.004)	0.5	(0.004)
Root mean squared error [f]	13.0	(0.036)	7.0	(0.022)	7.1	(0.023)	7.0	(0.022)
Confidence limit coverage [g]	8.0		94.2		94.1		94.2	

Abbreviations: IPW, inverse-probability weighting.

[a] Complete opioid information is available for 30% of the cohort.

[b] Defined as the difference between the estimated and true risk difference.

[c] Defined as the standard deviation of the bias across 10,000 simulation experiments.

[d] Defined as the standard error based on the standard deviation from 200 nonparametric bootstrap resamples averaged across 10,000 simulation experiments.

[e] Defined as the sum of the squared bias and variance.

[f] Defined as the square root of the sum of the squared bias and variance.

[g] Defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation trials.

Table 5.3b. Performance of g-computation approaches addressing missing data to estimate average intervention effect of shortening opioid prescription duration relative to the status quo among a cohort of 1,623 HIV-positive women, 10,000 simulation experiments. [a]

	Complete Case		IPW		Imputation	
	Estimate	Monte Carlo error	Estimate	Monte Carlo error	Estimate	Monte Carlo error
Bias [b]	7.4	(0.023)	0.5	(0.034)	0.5	(0.034)
Standard deviation of bias [c]	2.3		3.4		3.4	
Average standard error [d]	2.3	(0.002)	3.4	(0.004)	3.3	(0.003)
Mean squared error [e]	0.7	(0.004)	0.2	(0.002)	0.2	(0.002)
Root mean squared error [f]	7.8	(0.022)	4.7	(0.015)	4.5	(0.014)
Confidence limit coverage [g]	9.9		93.7		93.6	

Abbreviations: IPW, inverse-probability weighting.

[a] Complete opioid information is available for 30% of the cohort. Extrapolation approach is not shown because only the average treatment effect is estimated with this approach.

[b] Defined as the difference between the estimated and true risk difference.

[c] Defined as the standard deviation of the bias across 10,000 simulation experiments.

[d] Defined as the standard error based on the standard deviation from 200 nonparametric bootstrap resamples averaged across 10,000 simulation experiments.

[e] Defined as the sum of the squared bias and variance.

[f] Defined as the square root of the sum of the squared bias and variance.

[g] Defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation trials.

Table 5.4a. Performance of g-computation estimators of the average treatment effect under varied sample conditions with 30% complete opioid information.

Parameter	Null Intervention Effect [a]					Missing Completely at Random [b]				
	Truth	CC	Extrap	IPW	Impute	Truth	CC	Extrap	IPW	Impute
Risks										
>7 days	30.5	18.0	30.9	30.5	30.6	37.9	37.9	38.0	38.0	38.0
≤7 days	30.5	28.3	30.6	30.8	30.9	19.4	19.8	19.9	19.9	19.9
Risk difference										
≤7 vs >7 days	0.0	10.3	0.3	0.3	0.3	-18.6	-18.3	-18.1	-18.1	-18.1
Performance [e]										
Bias [f]		10.3	0.3	0.4	0.3		0.3	0.5	0.5	0.5
STD of bias [g]		3.9	5.6	5.6	5.6		4.4	4.2	4.2	4.2
Average SE [h]		3.9	5.6	5.6	5.6		4.3	4.1	4.1	4.1
MSE [i]		1.4	0.6	0.6	0.6		0.4	0.3	0.3	0.3
RMSE [j]		11.2	7.6	7.6	7.6		5.9	5.6	5.6	5.6
Coverage [k]		24.6	94.4	94.5	94.4		94.5	94.6	94.6	94.6

Table 5.4a (continued)

Parameter	Increased Effect Measure Modification [c]					Increased Sample Size [d]				
	Truth	CC	Extrap	IPW	Impute	Truth	CC	Extrap	IPW	Impute
Risks										
>7 days	37.9	23.9	37.9	37.9	37.9	37.9	23.9	37.9	37.9	37.9
≤7 days	19.7	22.7	19.9	19.9	19.9	19.4	17.7	19.5	19.5	19.5
Risk difference										
≤7 vs >7 days	-18.2	-1.3	-18.0	-18.0	-18.0	-18.5	-6.3	-18.4	-18.3	-18.4
Performance [e]										
Bias [f]		16.9	0.2	0.2	0.2		12.3	0.2	0.2	0.2
STD of bias [g]		1.8	2.3	2.6	2.3		1.8	2.5	2.6	2.5
Average SE [h]		1.8	2.3	2.6	2.3		1.7	2.4	2.6	2.4
MSE [i]		2.9	0.1	0.1	0.1		1.6	0.1	0.1	0.1
RMSE [j]		17.0	3.2	3.5	3.2		12.4	3.3	3.5	3.3
Coverage [k]		0.0	94.7	95.1	94.7		0.0	94.2	94.6	94.2

Abbreviations: CC, complete case; Extrap., extrapolation; IPW, inverse-probability weighting; STD, standard deviation; SE, standard error; MSE, mean squared error; RMSE, root mean squared error.

[a] Cohort of 1,623 HIV-positive women where shortening duration of opioid prescription has no effect on 12-month risk of emergency room/hospital visit or death.

[b] Cohort of 1,623 HIV-positive women where opioid information is missing completely at random.

[c] Cohort of 1,623 HIV-positive women where the presence of covariate Z (which also influences the probability of complete information and the outcome) modifies the relationship between opioid duration and the outcome in opposite directions; Z increases risk of outcome for those with long duration opioid prescriptions and decreases the outcome for those with short duration opioid prescriptions.

[d] Cohort of 7,500 HIV-positive women.

[e] Presented for estimates of the risk difference quantifying the effect of shortening opioid prescriptions among all women versus lengthening prescriptions among all women.

[f] Defined as the difference between the estimated and true risk difference.

[g] Defined as the standard deviation of the bias across 10,000 simulation experiments. .

[h] Defined as the standard error based on the deviation from 200 nonparametric bootstrap resamples averaged across 10,000 simulation experiments.

[i] Defined as the sum of the squared bias and variance.

[j] Defined as the square root of the sum of the squared bias and variance.

[k] Defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation trials.

Table 5.4b. Performance of g-computation estimators of the average intervention effect under varied sample conditions with 30% complete opioid information. [a]

Parameter	Null Intervention Effect [b]				Missing Completely at Random [c]			
	Truth	Complete Case	IPW	Imputation	Truth	Complete Case	IPW	Imputation
Risks								
≤7 days	30.5	28.3	30.8	30.9	19.4	19.8	19.9	19.9
Status quo	30.5	22.0	30.5	30.6	30.5	30.7	30.7	30.8
Risk differences								
≤7 vs status quo	0.0	6.3	0.3	0.3	-11.1	-10.9	-10.8	-10.8
Performance [f]								
Bias [g]		-8.4	0.1	0.1		0.2	0.3	0.3
STD of bias [h]		2.4	3.5	3.5		2.7	2.7	2.6
Average SE [i]		2.5	3.5	3.5		2.7	2.6	2.6
MSE [j]		0.1	0.2	0.2		0.1	0.1	0.1
RMSE [k]		6.8	4.7	4.7		3.7	3.6	3.5
Coverage [l]		27.4	95.0	94.9		94.6	94.7	94.6

Table 5.4b (continued)

Parameter	Increased Effect Measure Modification [d]				Increased Sample Size [e]			
	Truth	Complete Case	IPW	Imputation	Truth	Complete Case	IPW	Imputation
Risks								
≤7 days	19.7	22.7	19.0	19.0	19.4	17.7	19.5	19.5
Status quo	30.6	23.4	30.6	30.7	30.5	21.4	30.5	30.5
Risk differences								
≤7 vs status quo	-10.9	-0.7	-10.7	-10.7	-11.1	-3.8	-10.9	-11.0
Performance [f]								
Bias [g]		10.2	0.2	0.2		7.3	0.1	0.1
STD of bias [h]		1.1	1.8	1.6		1.1	1.8	1.6
Average SE [i]		1.1	1.8	1.5		1.1	1.7	1.6
MSE [j]		1.1	0.1	0.0		0.6	0.1	0.1
RMSE [k]		10.2	2.4	2.1		7.4	2.4	2.1
Coverage [l]		0.0	94.6	94.5		0.0	94.4	94.2

Abbreviations: IPW, inverse-probability weighting; STD, standard deviation; SE, standard error; MSE, mean squared error; RMSE, root mean squared error.
[a] Extrapolation approach is not shown because only the average treatment effect is estimated with this approach.
[b] Cohort of 1,623 HIV-positive women where shortening duration of opioid prescription has no effect on 12-month risk of emergency room/hospital visit or death.
[c] Cohort of 1,623 HIV-positive women where opioid information is missing completely at random.
[d] Cohort of 1,623 HIV-positive women where the presence of covariate Z (which also influences the probability of complete information and the outcome) modifies the relationship between opioid duration and the outcome in opposite directions; Z increases risk of outcome for those with long duration opioid prescriptions and decreases the outcome for those with short duration opioid prescriptions.
[e] Cohort of 7,500 HIV-positive women.
[f] Presented for estimates of the risk difference quantifying the effect of shortening opioid prescriptions relative to the status quo.
[g] Defined as the difference between the estimated and true risk difference.
[h] Defined as the standard deviation of the bias across 10,000 simulation experiments. .
[i] Defined as the standard error based on the deviation from 200 nonparametric bootstrap resamples averaged across 10,000 simulation experiments.
[j] Defined as the sum of the squared bias and variance.
[k] Defined as the square root of the sum of the squared bias and variance.
[l] Defined as the proportion of times the 95% confidence limit trapped the true parameter value across the 10,000 simulation trials.

CHAPTER 6: DISCUSSION

A. Overview of Key Findings

The overall objectives of this work were to 1) use novel quantitative methods to estimate the long-term impact of combined interventions on non-AIDS risk factors among women with HIV and 2) develop two-stage approaches to estimate treatment and intervention effects from observational HIV data when exposure information is partially missing.

A1. Aim 1

In Aim 1 we were able to use the parametric g-computation algorithm and nearly two decades of observational data collected in the WIHS to estimate the risks of all-cause mortality that would be observed under combined interventions on non-AIDS risk factors. Much attention to date has focused on evaluating ART-based interventions without considering intervention portfolios that more comprehensively address the complex health threats experienced by those with HIV. Yet, ART treated individuals with HIV still face greater health disparities and a higher burden of comorbidities than the general population.[1–4] With 65% of those with HIV estimated to be affected by 2 or more chronic conditions,[16] morbidity solely attributable to HIV is the exception, not the norm. Thus, evaluating the long-term impacts of interventions on non-AIDS risk factors is critical to the development of improved healthcare delivery guidelines among those with HIV.

We estimated the risks of mortality in the WIHS under interventions on alcohol and smoking combined with prompt initiation of ART in the modern treatment era. We estimated risks under interventions (nearly) eliminating each of these risk factors as well as risks under more realistic interventions that reduced the prevalence of these risk factors based on the expected efficacy of existing interventions. To quantify the impact of additionally targeting non-AIDS risk factors compared to status

quo healthcare delivery among those with HIV, we compared the risk of mortality under each intervention portfolio with the risk of mortality under an intervention on prompt initiation of modern ART alone using risk differences and ratios.

Our results suggest that while the risk of all-cause mortality in the WIHS dramatically decreases with prompt initiation of modern ART, further reductions in the mortality risk are possible with interventions on non-AIDS risk factors. We specifically evaluated the effects of interventions on alcohol and smoking because previous research has shown more pronounced adverse effects of these factors among those with HIV.[15,41–47] Additionally, detailed information on alcohol and smoking has been systematically collected in the WIHS since its outset which offers an ideal opportunity for leveraging rich, longitudinal data from a generalizable study population to evaluate potential intervention effects.

The 8-year risk of all-cause mortality was lowest under interventions on smoking as opposed to interventions on alcohol. This was unsurprising given that the prevalence of smoking in the WIHS is much higher than that of alcohol use. Additionally, whereas the adverse effects of smoking on morbidity and mortality among populations with HIV are strong,[41] the effects of various levels of alcohol consumption among those with HIV, particularly women, are not as well-understood. It is possible that the effects of alcohol consumption levels among women with HIV varies by characteristics other than HCV status.

While all sets of intervention portfolios suggested improved survival in the WIHS, we found that sets that eliminated rather than reduced the prevalence of non-AIDS risk factors were most effective. This was particularly evident for the intervention on smoking. The intervention that combined universal initiation of modern ART with elimination of smoking (which assumes an intervention with 100% efficacy) reduced the 8-year risk of mortality by 18% compared to an intervention on modern ART initiation alone. However, the intervention that combined universal initiation of modern ART with a real-world behavioral smoking cessation intervention reduced the 8-year risk of mortality by 9%. The main reason for this difference is that the existing intervention has only a 0.16 probability of achieving smoking

cessation. Yet, this is more representative of the expected efficacy of many other existing smoking cessation interventions.

Overall, our findings from Aim 1 suggest that noteworthy improvements in survival may be achieved with interventions targeting alcohol and smoking risk behaviors among women with HIV. To obtain the best outcomes, however, more efficacious interventions are needed. Additionally, interventions targeting other non-AIDS risk factors should be considered and may more significantly reduce the risk of all-cause mortality as compared to interventions on alcohol.

A2. Aim 2

In Aim 2 we used missing data theory and two-stage study design approaches to develop two-stage g-computation estimators of the average treatment and intervention effects that can be used in the presence of missing exposure information. With attention shifting to more comprehensive healthcare delivery strategies among those with HIV, existing observational cohorts will continue to be an important source of information for generalizable and cost-effective studies of interventions on non-AIDS risk factors. For example, it is likely that future studies will entail intervention trials nested in existing cohorts or enhancing previously collected cohort data with data from other sources to form hybrid study designs. While the parametric g-computation algorithm is a powerful tool in using observational data to estimate intervention effects in a population of interest, it requires adaptations to accommodate challenges that will arise with missing and imperfectly measured data in these situations.

We developed and validated an extrapolation g-computation estimator of the average treatment effect and two-stage inverse probability weighted and exposure imputation g-computation estimators of the average intervention effect in a simulation study motivated by the WIHS. We considered a hypothetical scenario in which the goal was to estimate the effect of opioid prescription duration on the 12-month risk of hospital/emergency room visit or death in the WIHS, supplementing existing WIHS data with electronic medical record linkages to obtain prescription information for a sample of WIHS participants in regular HIV care. To illustrate and validate the two-stage approaches, we applied estimators in a cohort simulated to represent 1623 women in the WIHS in 10,000 Monte Carlo simulation

trials. We varied the percentage of missing opioid prescription information from 10-90% and examined additional scenarios in which there was a null treatment effect, data were missing completely at random, there was increased effect measure modification, and there was an increased sample size. To assess operating characteristics of approaches, we estimated bias, average standard error, root mean squared error, and confidence limit coverage.

Our results suggest that all two-stage estimators explored have good operating characteristics in the presence of missing exposure information compared to a naïve g-computation estimator fit to the sample with complete data. The two-stage extrapolation g-computation estimator of the average treatment effect was unbiased across all scenarios. While the average standard error was higher than that of the complete case analysis, the reduction in bias resulted in a smaller root mean squared error. The two-stage extrapolation estimator also approached the nominal 95% confidence limit coverage. Similarly, the two-stage inverse probability and exposure imputation g-computation estimators of the average intervention effect were unbiased, had increased average standard error but decreased root mean squared error, and approached nominal 95% confidence limit coverage.

Overall, our findings in Aim 2 demonstrate that naïve g-computation estimators fit to complete cases can lead to significantly biased results in the presence of missing exposure information but two-stage g-computation estimators can overcome this limitation. It is possible to leverage complete covariate and outcome data on the total study population of interest along with partially observed exposure information to estimate average treatment and intervention effects. This presents new avenues for cost-effective, efficient study designs of non-AIDS-related interventions necessary to answer pressing questions regarding optimal healthcare delivery among those with HIV in the modern ART era.

B. Limitations

There are important limitations of our work. In Aim 1, these include imperfect measures and variable definitions as well as potential violations in modeling assumptions. Aim 2 was primarily limited by the structure of our simulation design.

First, it should be noted that when estimating the effects of interventions on non-AIDS risk factors in the WIHS in Aim 1, we used self-reported alcohol and smoking information. Alcohol consumption and smoking status is reported by WIHS participants during semi-annual study visits and may be subject to measurement error. Due to social desirability bias, it is possible that WIHS participants may underreport alcohol and smoking. This may lead to underestimating alcohol use and smoking in the WIHS in which case we would have intervened on fewer women than should have actually received interventions based on their true risk factor status. However, in practice, it is the potentially mismeasured self-reported risk factor information that clinicians will likely use to make intervention decisions.

Second, there were three main limitations specific to our interventions on alcohol use. Among those with HCV, alcohol interventions were based on HCV status at baseline. Clinicians might be more likely to use HCV status over time to make decisions about alcohol consumption at each time point. We used baseline measures of HCV, because data were available for most WIHS participants at this visit. However, future studies will likely benefit from accounting for HCV clearance or new infections.

Additionally, our specified interventions on alcohol consumption among those without HCV relied on a somewhat arbitrary limit of alcohol consumption. National recommended safe limits for women are no more than 7 drinks per week and no more than 3 drinks in one day.[119] In men with HIV, however, lower thresholds of alcohol consumption result in higher risk of physiologic injury and mortality than those without HIV.[47] Analogous studies are absent for women, but we considered that they may similarly be adversely affected by alcohol at lower thresholds of consumption than those dictated by national recommendations for the general population. In sensitivity analyses, we assessed alternate cut points but more restrictive limits did not meaningfully change results. Future research will be needed to define appropriate alcohol consumption limits for both men and women with HIV.

Finally, the interventions we modeled to reduce the prevalence of alcohol consumption were less well-defined than the intervention on smoking. We used the expected efficacy of an alcohol reduction intervention among those with HCV, but these results were from a small sample of patients at HCV clinics and was not specific to women with HIV.[120] Because alcohol interventions are typically conducted

among those with more severe alcohol dependency, there was little information regarding the efficacy of interventions to lower drinking among those without as extreme consumption levels. Thus, for those without HCV, we modeled a hypothetical intervention in which it was possible to limit drinking to 3 drinks per week among all women reporting higher levels of consumption.

Third, our definition of prompt initiation of modern ART was broader than would have been ideal. Because WIHS is an interval cohort and visits occur semi-annually, we needed to define "prompt" initiation as initiation occurring by the first follow-up visit. This is a crude measure of time to ART initiation, but we could not further distinguish time to initiation because initiation dates in the WIHS are recorded as the first date observed to be on ART while under WIHS study. Currently, research is focusing on "immediate" or "rapid" ART initiation strategies in which timing of ART initiation occurs within days or weeks of HIV diagnosis.[153–155] This is likely a more relevant window and would have been considered had data been available. Furthermore, we defined "modern ART" as initiation of ART at/after 1 October 2001 (the period after which tenofovir had been approved and became a common drug in combination ART regimens).[114] This is analogous to an intent-to-treat design in which we assume individuals initiating ART in this time period initiated modern ART regimens. We did not specifically assess whether regimens contained tenofovir, so it is possible that not all women initiating ART during this period initiated modern regimens.

Fourth, we assumed a set of sufficient conditions for parametric g-computation estimation of the effect of interventions on mortality in the WIHS which may have been violated to some extent. We assumed consistency which is often described as no interference and no treatment version relevance.[108,109] Consistency entails that the risk of mortality for WIHS participants who were observed to receive the intervention of interest is equal to the risk of mortality that would be observed if the intervention status was set to the value of interest. For example, the risk of mortality for women who were observed to promptly initiate ART in the modern treatment era is equal to the risk of mortality for women if they had not "naturally" promptly initiated ART in the modern treatment era but, rather, were set to do so.

Interference is not a concern in this setting as it is unlikely a woman's intervention portfolio (e.g., receiving the smoking cessation intervention combined with modern ART initiation) would affect another woman's risk of mortality. In evaluating the potential for treatment version relevance, we need to consider whether varying versions of the intervention have differing effects on mediating variables and outcomes. In regards to prompt initiation of modern ART, there might be some treatment version relevance that is masked by our crude definition of ART timing. For example, initiating ART within 1 month versus 5 months of study entry likely has differing impacts on time-varying covariates (e.g., CD4 count and viral load) and outcomes. Because we cannot distinguish between these timing strategies, however, our intervention represents an ART initiation time that is averaged over unobserved initiation times. This may limit generalizability to a setting in which the distribution of unobserved initiation times is different.

With smoking and alcohol interventions, assessing consistency is more complicated because we used efficacy estimates from outside of the WIHS. Thus, we needed to make a further assumption that there are no indirect pathways between the intervention and outcome that do not pass through the exposure[52] (i.e., the intervention exclusion restriction assumption); or, if there are indirect pathways, they are appropriately modeled. For the smoking intervention, we assumed that the behavioral intervention did not have effects on variables other than smoking status. For the alcohol intervention, we assumed that the intervention did not have effects on variables other than alcohol intake (though we did allow changes in alcohol intake to effect smoking status). However, we tried to select interventions that had minimal effects on indirect pathways to mortality. Smoking and alcohol interventions that include more intensive cognitive behavioral therapies or pharmacotherapies were likely to have greater potential for violations of the intervention exclusion restriction assumption.

In using the parametric g-computation algorithm we also assumed conditional exchangeability (i.e., no unmeasured confounding or selection bias).[110] Unmeasured confounding is always a threat to observational studies. In our setting, there was likely a greater threat to conditional exchangeability for smoking and alcohol interventions than for ART initiation. Because guidelines for smoking and

especially alcohol reduction interventions are less well-defined than those for ART initiation in HIV populations, it is possible uncontrolled confounding remains. However, the rich data collected by the WIHS allowed us to address key baseline and time-varying confounders that would not otherwise have been possible to account for in clinical settings. In evaluating the potential for selection bias, it is important to consider those who were eligible for our study analysis but were excluded for missing baseline information. As those excluded represented only 2% of the eligible sample, any differences between those included and excluded likely have minimal impacts on results.

We also assumed no measurement error,[107] positivity,[111] and correct specification of parametric models. Potential issues with measurement error were described above in regards to self-reported alcohol and smoking. To improve positivity (i.e., nonzero probability of exposures within all strata of covariates) we carefully selected covariates, functional form, and interaction terms to avoid cells with sparse data. While it is not possible to verify correct model specification, we compared the data distribution simulated by g-computation to that observed in the WIHS cohort and confirmed we were able to replicate the observed distributions.

In Aim 2, our main limitation was that our simulation was conducted in a limited region of the possible parameter space (as is the nature of simulation studies). Therefore, our findings might not generalize to all research settings. However, we considered a number of scenarios and clearly reported the conditions under which simulations were performed.

Additionally, we considered a simplified scenario in which exposure was time-fixed, there were not multiple points of study follow-up, and data were only missing for the exposure as opposed to both exposure and covariate or outcome data. Because data were only missing for exposure, this also reduced our scenario to an example of monotone missing data. Nonmonotonic missing data patterns are likely to be encountered in HIV study settings and require nuanced approaches particularly when using inverse probability weighting approaches to address missing information.[156,157] Furthermore, in longitudinal settings with missing exposure and/or covariate data at several timepoints, approaches to addressing missing data become more complex. However, our work took the essential first step in addressing missing

data within the parametric g-computation framework and will provide the foundation for further extensions to these increasingly complex scenarios.

C. Strengths

Our work makes an important contribution to the literature as one of the only studies to evaluate the long-term effects of combined interventions on non-AIDS risk factors among people with HIV as well as the first to develop two-stage g-computation estimators that can be applied to estimate such effects when exposure information is partially missing. With few exceptions,[64,65] the implementation of the parametric g-computation algorithm in HIV settings has been restricted to the estimation of an intervention on one risk factor at a time. In particular, most previous applications have focused primarily on AIDS-related risk factors (i.e., ART treatment and timing of initiation).[68,73,75] This is likely due to the fact that AIDS-related comorbidities used to account for the majority of early mortality among those with HIV.[158–160] Therefore, studies of effective HIV treatment and optimal timing of ART initiation were the priority. Because ART has typically been a well-measured exposure captured throughout the last two decades of clinical and observational cohort follow-up, this has presented an ideal scenario for evaluating effects of ART-based interventions using the g-computation algorithm.

Yet, in the modern ART treatment era, it is now becoming clear that interventions on ART initiation alone have not closed the gap in survival for those with HIV who continue to die from non-AIDS causes 8 to 9 years earlier than those without HIV.[20–22] Given the complex clustering of multiple comorbidities, it is necessary to evaluate and compare the long-term impacts of various interventions on non-AIDS risk factors to prioritize strategies and ultimately reduce persisting disparities. Additionally, cost-effective and efficient evaluation of comprehensive interventions using existing cohorts requires adaptations to the g-computation algorithm that can accommodate frequently missing and imperfectly measured information on non-AIDS risk factors.

Our work addressed both of the points above. First, we were able to use the g-computation algorithm to evaluate and compare the long-term effects of interventions on alcohol and smoking. Consequently, our study is one of the first to provide estimates of the potential long-term impacts that

may be achieved with more comprehensive intervention portfolios. Second, we developed two-stage g-computation estimators that can be used to estimate effects of interventions on risk factors that are less well-measured or systematically collected in existing cohorts. Thus, our study was the first to provide investigators with validated, sophisticated approaches to implement g-computation methods in the context of missing data without sacrificing efficiency or validity.

Use of the WIHS for Aim 1 provided several noteworthy benefits. WIHS is the largest and longest ongoing interval cohort study of women with HIV in the US. This provided us with a plethora of longitudinal data that have been systematically collected at semi-annual study visits since 1998. In particular, the WIHS has collected rich behavioral data that are not always available in clinical cohorts. The length of follow-up time allowed us to evaluate impacts of interventions on long-term outcomes (i.e., 8-year risk of mortality) rather than suboptimal short-term proxies as is often the case in randomized trial settings. Finally, WIHS's focus on women is especially important, given that most US-based HIV clinical cohorts include over 75% men[63,161] and randomized controlled trials frequently exclude or limit enrollment of women; yet, it is women who experience the greatest burden of multimorbidity among those with HIV.[5] Thus, WIHS is a critical source of information for studies examining multimorbidity among those with HIV and providing results that can reduce disparities in comorbidities among women. Furthermore, women enrolled in the WIHS have a race/ethnicity distribution representative of women with HIV in the US which improves the generalizability of our findings.[91]

Another advantage of our work in Aim 1 is that we used prompt initiation of ART in the modern treatment era as our reference for comparing portfolios combining ART with interventions on non-AIDS risk factors. We chose this reference as opposed to the natural ART initiation course or prompt initiation of ART which could have occurred before or after the availability of modern regimens because it is a timelier, clinically relevant comparison.[51] Our research question was based on determining the additional impact on survival that could be achieved by upscaling alcohol and smoking interventions compared to the status quo healthcare delivery among those with HIV. Since the beginning of WIHS follow-up, ART initiation guidelines have changed both in terms of timing of ART initiation and in regimen availability.

Thus, we were more interested in contrasting interventions on non-AIDS risk factors with an intervention that is founded in current HIV care guidelines – guidelines that focus on getting everyone on modern ART regimens without delay. Effect contrasts (e.g., risk differences and the related "number needed to intervene") that leverage this reference are important to clinicians in making decisions and policymakers for cost-effectiveness analyses.[49–52]

Our estimation of intervention portfolios both eliminating and reducing the prevalence of non-AIDS risk factors is an innovative and highly significant approach. Many applications of the parametric g-computation algorithm model hypothetical interventions that completely eliminate a risk factor or set a risk factor to some threshold level among 100% of the study sample of interest.[51,69,162] While this can be informative of the lower bound in mortality risk that could be achieved with perfectly efficacious interventions (assuming no other effects of interventions on pathways to the outcome aside from those modeled through exposure),[46,52] its policy relevance is limited.[49–51] Perfectly efficacious interventions do not exist in most cases. Thus, estimated effects of such hypothetical interventions do not correspond to changes in outcomes that are likely to be observed when real-world interventions are implemented. This makes it extremely difficult for clinicians to determine which risk factors are most important to target with existing interventions to achieve optimal outcomes as well as for policymakers to assess the cost-effectiveness of various interventions.

Westreich recently highlighted these challenges,[46,49,50] arguing for a new vocabulary for causal contrasts with greater relevance to public health policy.[51] He outlined several intervention contrasts that are critical for implementation science, including the generalized intervention contrast in which outcomes in the population are contrasted with outcomes under a scenario in which exposure is reduced but not eliminated.[51] This takes into account the fact that not everyone eligible for intervention may be reachable, not everyone reachable may be willing to receive intervention, and importantly, that for all those receiving intervention, the intervention will not perfectly eliminate the relevant risk factor. Thus, generalized intervention contrasts provide a more realistic view of how changes in policy and clinical guidelines may improve population-level outcomes in real-world settings.

Our work contributes to the limited research to date that has provided estimates of generalized intervention effects from observational data. In other settings, this has often been performed by using reported efficacy estimates from trial data or meta-analyses to model reductions in the prevalence of risk factors that are likely to be achieved by the given intervention.[46,64,65,103] We used estimates from a behavioral smoking cessation intervention implemented in a trial setting as well as an alcohol reduction intervention implemented in an HCV clinic to model expected reductions in the prevalence of smoking and alcohol that would be achieved with these interventions. These reductions of 16% and 36-62% for smoking and alcohol interventions respectively were much smaller than the 100% efficacy assumed by our elimination strategies. Consequently, while risks of mortality with interventions on non-AIDS risk factors still suggested improved survival, improvements were much smaller than those under complete elimination of risk factors (particularly for smoking). However, it is this information on which future clinical guidelines and policy decisions relies. Our estimates of the mortality risk under both elimination and realistic reduction of alcohol and smoking demonstrate that elimination of these risk factors can improve survival but existing interventions are not efficacious enough to achieve these improvements. Further research will be needed to identify better smoking cessation and alcohol elimination interventions.

Part of this further research will likely entail leveraging existing cohorts to assess additional alcohol and smoking intervention effects and effects of interventions on other non-AIDS risk factors using the parametric g-computation algorithm. A remaining barrier that needs to be confronted, however, is frequently missing and imperfectly measured exposure information – both in existing cohorts as well as hybrid study designs in which existing cohort data may be supplemented with other data sources. Methods for accounting for missing data in the parametric g-computation framework have not been outlined or validated. Thus, the validity of findings from parametric g-computation applications using ad-hoc methods to account for missing data in these settings is questionable.

A major strength of our work in Aim 2 is that we adapted existing missing data theory and two-stage analytic approaches for the parametric g-computation algorithm and validated these methods for application in future studies. Our use and validation of existing methods applied to the parametric g-

computation algorithm allows investigators to easily construct our developed estimators in common statistical software programs for epidemiologic analyses. Therefore, these approaches are accessible and immediately available to all those who may want to use them.

We also based our simulation example in a realistic setting and used the WIHS to inform parameters. Thus, results have improved generalizability to the research scenarios investigators are likely to encounter. The wide range of scenarios we assessed under varying sample conditions provides a comprehensive picture of our estimators' operating characteristics.

D. Summary and Public Health Significance

Despite over two decades of effective ART, people with HIV experience persisting disparities in morbidity and mortality compared to the general US population, dying from non-AIDS causes 8 to 9 years earlier than those without HIV.[20,21] Our work presents some of the first results regarding the potential improvement in survival that may be achieved with comprehensive healthcare delivery strategies that additionally target non-AIDS risk factors prevalent in populations with HIV. Given the complex clustering of multiple comorbidities, especially among women with HIV,[5] it is necessary to evaluate and compare the long-term impacts of various interventions on non-AIDS risk factors to prioritize strategies and ultimately improve healthcare delivery.

We have illustrated how observational HIV cohort data and novel quantitative methods can be leveraged to evaluate long-term impacts of interventions that would be difficult to assess in randomized controlled trial settings. While randomized trials were essential in informing current ART treatment policies including regimen and timing of ART initiation,[163,164] important limitations of this study design have impeded its use in assessing and comparing the impact of interventions on non-AIDS risk factors. These include (1) the large sample size required to conduct factorial randomized trials to assess multiple, combined interventions, (2) the length of follow-up required to evaluate long-term effects on clinical outcomes of interest rather than on their imperfect, short-term proxies, and (3) stringent inclusion and exclusion criteria which limit generalizability of results – especially for women, racial and ethnic minorities, and those with complex health histories who are most in need of intervention. Observational

HIV cohorts, on the other hand, have followed large numbers of HIV-positive individuals in the US since the early HIV epidemic. Yet, confounding remains a major concern when using traditional study approaches in observational cohorts to quantify effects of treatments or interventions due to nonrandomization of treatment.[135]

Recent illustrations of the parametric g-computation algorithm[68,69,71–73,75] now offer opportunities to leverage observational data to quantify the impacts of interventions on non-AIDS risk factors, provided that relevant exposures have been well-measured and documented over the course of cohort follow-up. In Aim 1, we demonstrated such an application in an innovative manner that considered multiple intervention portfolios and generalized intervention contrasts.[46,50,51] Our work provides investigators with a framework for conducting comparative effectiveness studies in HIV cohorts as well an example of how one can estimate effects of real-world interventions with imperfect efficacy using observational data. This will be critical to addressing health disparities and improving health outcomes among ART-treated women with HIV in the current treatment era. Additionally, our finding that reductions in the risks of all-cause mortality were attenuated under real-world interventions on alcohol and smoking as compared to the hypothetical elimination of these risk factors calls for identification of more efficacious interventions. Our work illustrates both the potential improvement in survival that can be achieved with perfectly efficacious interventions and the significant gap of existing interventions in terms of approaching this 100% efficacy.

While the parametric g-computation algorithm approach implemented in Aim 1 can be used by investigators to further evaluate interventions on non-AIDS risk factors among those with HIV when exposures have been well-measured and documented over the course of cohort follow-up, missing data are extensive in epidemiologic studies.[77,165] Complete case analyses can lead to imprecise and biased results[79,80,131] and presents a threat to the validity of results obtained when leveraging this powerful analytic tool in frequently encountered settings of incomplete exposure information. Our work in Aim 2 has provided investigators with a framework for leveraging limited exposure information available on a subset of the study population in two-stage g-computation estimation. Our developed estimators provide

investigators with improved approaches to implement causal inference methods in the context of missing data without sacrificing efficiency or validity. Further, these approaches can guide more efficient study design when resources are limited especially when considering nested trials and hybrid study designs in which existing cohort data may be supplemented with other data sources.

In conclusion, this body of work provides a framework for producing clinical and policy-relevant comparative estimates of the impact of potential comprehensive healthcare delivery strategies among people with HIV using existing, imperfect longitudinal data sources. Our illustration, development, and assessment of statistical approaches for leveraging observational data in realistic research settings will facilitate the broader application of these methods to answer important public health questions among HIV-positive and other populations. This can also improve the development of evidence-based interventions in a timely manner, using the information presently available. Finally, our outlined approaches can help researchers design more accurate studies in the future.

APPENDIX A: SUPPLEMENTAL TEXT

Overview of the Parametric G-computation Algorithm

We estimated the cumulative incidence of all-cause mortality in the Women's Interagency HIV Study (WIHS)[89–91] under each set of intervention portfolios (i.e., intervention combinations) using the parametric g-computation algorithm.[67–69,104,105] Analogous to inverse probability of treatment weighting,[76,100,135] this approach is a generalization of standardization that may be applied to observational data to re-express the outcome distribution under observed exposure as the outcome distribution that would be observed under changes in, or interventions on, exposure. In longitudinal settings subject to time-varying confounding, the parametric g-computation algorithm can provide valid estimates of the exposure effect that would otherwise be biased with the use of traditional regression models.[67,101]

Applications of the parametric g-computation algorithm to estimate the risk of mortality under interventions on antiretroviral therapy (ART) initiation using data from observational HIV cohorts have yielded results comparable to those found in randomized controlled trial settings.[68,75] Three noteworthy advantages of this estimator are its flexibility in estimating population-level effects of realistic changes to exposure distributions,[46,49–51] ability to estimate effects of interventions that depend on the natural value of exposure,[74,104,105] and efficiency in quantifying impacts of multiple interventions.[69,105] Detailed descriptions of the parametric g-computation and illustrations of its implementation are available.[46,65,162,68–70,74,75,97–99] Below, we describe the implementation of this approach in our research context.

Notation

Let uppercase letters denote random variables and lowercase letters denote potential realizations of those variables. Let i index 1, …, 1016 women in our study sample, j index 1, …, J completed visits of follow-up, and Y_{ij} represent an indicator of death for woman i at time j. The maximum number of follow-up visits is $J = 16$; as WIHS visits occur every 6 months, this corresponds to 8 years of follow-up from study baseline. We use A_{ij} to represent a binary indicator of antiretroviral therapy (ART) treatment,

D_{ij} to represent categorical indicators for average drinking level per week and S_{ij} to represent a binary indicator of smoking. We use $\boldsymbol{Z}_{ij}$ to represent a vector of time-varying covariates for woman i measured at time j (i.e., CD4 count, detectable viral load, and depression) and $\boldsymbol{W}_{ij}$ to represent a vector of time-fixed demographic and clinical characteristics for woman i measured at baseline (i.e., $j = 0$). The vector $\boldsymbol{W}_{ij}$ includes baseline age, race, level of education, history of injection drug use (IDU), prior exposure to dual therapy zidovudine (AZT), wave of WIHS enrollment, CD4 count, viral load, and smoking status. Finally, we use C_{ij} to represent a binary indicator of being censored due to loss to follow-up which we define as having two consecutive missed WIHS visits. All women are administratively censored after a maximum of 8 years of completed follow-up due to sparse data at later time points.

At each time point j, the temporal ordering of variables is: $\boldsymbol{W}_{ij}$; C_{ij}; Y_{ij}; A_{ij}, D_{ij}, S_{ij}; $\boldsymbol{Z}_{ij}$ (Supplemental Figure 1). The values of the time-fixed covariates $\boldsymbol{W}_{ij}$ remain constant for each woman throughout the study period so that for a given woman i, $\boldsymbol{w}_{i0} = \boldsymbol{w}_{i1} = \ldots = \boldsymbol{w}_{ij}$. Among those who remain uncensored and alive by visit j (i.e., $C_{ij} = Y_{ij} = 0$), values of ART (i.e., A_{ij}), drinking level (i.e., D_{ij}), and smoking (i.e., S_{ij}) are based on self-report corresponding to the 6-month period between $j - 1$ and j.

It is important to note that due to the nature of the WIHS as an interval cohort, exposure to ART, alcohol, and smoking, while occurring between $j - 1$ and j, are documented at time j. It is unknown from the observed data whether, among those who die before a given follow-up visit, there was a change in exposure status in the time period between the last follow-up visit at which they remained alive (and uncensored) and their death date. ART initiation dates are recorded as the first WIHS visit date at which individuals report having initiated ART, while death dates are recorded as actual dates of death. Therefore, we assume that those who had not initiated ART at the last visit at which they remained alive and uncensored did not do so before their death. Similarly, we assume drinking and smoking status did not change between the last visit at which women remained alive and uncensored and their death. Or rather, in both cases, we assume that any change in ART, drinking, or smoking between the last follow-up

visit and a death occurring less than 6 months later was too short of an exposure period to notably affect the risk of death.

Among those who remain alive and uncensored at time j, measured values of time-varying covariates (i.e., $\boldsymbol{Z}_{ij}$) correspond to real time, or nearly real time, measures at j. Specifically, CD4 cell count and viral load are measured by laboratory tests conducted at the time of the WIHS visit. Depression is measured by a validated questionnaire[124] that asks women to report symptoms experienced during the last week. In the remaining text we suppress subscript i.

The cumulative incidence of mortality in the WIHS under the observed exposure history (i.e., no intervention on ART initiation or non-AIDS risk factors) at time j can be expressed as Equation 1:

$$F(j) = \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\bar{a}_j}\sum_{\bar{\boldsymbol{w}}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_k=\bar{a}_k,\bar{D}_k=\bar{d}_k,\bar{S}_k=\bar{s}_k,\ \overline{\boldsymbol{Z}}_k=\bar{\boldsymbol{z}}_k,\boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\ \bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\ \overline{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(S_m=s_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\overline{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(D_m=d_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\ \overline{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(A_m=a_m\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\overline{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}=\bar{a}_{m-1},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\overline{\boldsymbol{Z}}_{m-1}=\bar{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m)\times \end{bmatrix}\right\}$$

Parameters of Interest

For our base scenario (i.e., Portfolio 1), our parameter of interest is the cumulative incidence of mortality in the WIHS under universal, prompt initiation of ART in the modern treatment era. We adapt the above formula to reflect this parameter by setting $A = 1$, indicating that all women are being set to initiate ART in the modern treatment era. Specifically, this indicator represents having initiated therapy by the time of the first follow-up visit (i.e., $A = 0$ for all women at $j = 0$ and $A = 1$ for all women at $j = 1$ and all visits thereafter). This is akin to an intent-to-treat analysis in which we assume that, once therapy is initiated, it is continued throughout the study period. It should be noted that women enrolled in the WIHS prior to 1 October 2001 and set to initiate ART by the first follow-up visit could not have initiated modern ART. This is based on our definition of modern ART initiation as initiation of ART on, or after, 1 October 2001 (the period after which tenofovir had been approved and became a common drug in combination ART regimens).[114] However, we use observed exposure information from those in the

modern treatment era to extrapolate modern ART to this earlier time period. Our adapted incidence function is written as Equation 2:

$$F^{a^g}(j) = \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\boldsymbol{w}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_{k-1}^{g}=\bar{a}_k^g,\bar{D}_k=\bar{d}_k,\bar{S}_k=\bar{s}_k,\ \boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(S_m=s_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P\left(D_m=d_m\middle|\bar{D}_{m-1}=\bar{d}_{m-1},\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ 1\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}^{g}=\bar{a}_{m-1}^{g},\bar{D}_{m-1}=\bar{d}_{m-1},\bar{S}_{m-1}=\bar{s}_{m-1},\overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_m=\boldsymbol{w}_m,\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m) \end{bmatrix}\right\}$$

where we use superscript g to convey that the value of A is being set according to the intervention plan.

Our other parameters of interest are the cumulative incidences of mortality in the WIHS under universal, prompt initiation of ART in the modern treatment era combined with various interventions on alcohol consumption and/or smoking. Because we are interested in interventions that depend on the natural value of alcohol consumption and smoking (i.e., interventions that are implemented based on the values of alcohol and smoking that would be observed if intervention were discontinued immediately before measurement at time *j*), we use the extended version of the parametric g-computation algorithm.[104,105] The incidence function under these additional interventions can be expressed with the extended parametric g-formula as Equation 3:

$$F^{(a,d,s)^g}(j)$$

$$= \sum_{\bar{\boldsymbol{z}}_j}\sum_{\bar{s}_j}\sum_{\bar{d}_j}\sum_{\boldsymbol{w}_j}\sum_{k=0}^{j}\left\{\prod_{m=0}^{k}\begin{bmatrix} P\left(Y_{k+1}=1\middle|\bar{A}_k^g=\bar{a}_k^g,\bar{D}_k^g=\bar{d}_k^g,\bar{S}_k^g=\bar{s}_k^g,\ \overline{\boldsymbol{Z}}_k=\overline{\boldsymbol{z}}_k,\boldsymbol{W}_k=\boldsymbol{w}_k,\bar{C}_k=\bar{Y}_k=0\right)\times \\ f\left(\boldsymbol{Z}_m\middle|\bar{A}_{m-1}^g=\bar{a}_{m-1}^g,\ \bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\ \bar{S}_{m-1}^g=\bar{s}_{m-1}^g,\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^g\left(S_m^g=s_m^g\middle|S_m^*=s_m^*,\bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\ \bar{S}_m^g=\bar{s}_m^g,\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{obs}\left(S_m^*=s_m^*\middle|\bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\bar{S}_{m-1}^g=\bar{s}_{m-1}^g,\overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^g\left(D_m^g=d_m^g\middle|D_m^*=d_m^*,\ \bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ P^{obs}\left(D_m^*=d_m^*\middle|\bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_m=\ \bar{Y}_m=0\right)\times \\ 1\times \\ P\left(Y_m=0\middle|\bar{A}_{m-1}^g=\bar{a}_{m-1}^g,\ \bar{D}_{m-1}^g=\bar{d}_{m-1}^g,\ \bar{S}_{m-1}^g=\bar{s}_{m-1}^g,\ \overline{\boldsymbol{Z}}_{m-1}=\overline{\boldsymbol{z}}_{m-1},\boldsymbol{W}_{m-1}=\boldsymbol{w}_{m-1},\bar{C}_{m-1}=\ \bar{Y}_{m-1}=0\right)\times \\ f(\boldsymbol{W}_m) \end{bmatrix}\right\}$$

We use superscript g to convey that the values of D and S (i.e., alcohol consumption and smoking) are being set according to the intervention plan while we use an asterisk to denote the values that would be observed if the intervention plan were discontinued immediately before measurement of these variables. The value $\bar{a}_j^g$ is set to 1 across all $j \geq 1$ for each set of intervention portfolios while the values of $\bar{d}_j^g$ and $\bar{s}_j^g$ are set according to the plans below.

Set A Intervention Portfolios to Nearly Eliminate Non-AIDS Risk Factors:

Portfolio 2A: Under this portfolio, if observed drinking level is greater than zero for those with an indication of hepatitis C virus at baseline (i.e., positive antibody and positive or missing RNA status), drinking level is set to zero with probability 1. For those without an indication of hepatitis C virus, if observed alcohol consumption is above 1 drink per week, alcohol consumption is set to a limit of 1 drink per week. Otherwise, alcohol consumption is set equal to that of observed consumption.

Portfolio 3A: Under this portfolio, all women reporting smoking are set to quit smoking by the next follow-up visit.

Portfolio 4A: Under this portfolio, both Portfolio 2A and 3A are implemented.

Set B Intervention Portfolios to Reduce Non-AIDS Risk Factors:

Portfolio 2B: Among those with an indication of hepatitis C virus at baseline, drinking level is set to zero with probability 0.36, reduced by 50% with probability 0.26, and set to the observed value with probability 0.38 based on the reported efficacy of a brief alcohol counseling intervention.[120] Among those without an indication of hepatitis C virus, drinking level is set to a limit of 3 drinks per week among those observed to drink above this limit.

Portfolio 3B: Among those reporting smoking at a given visit, smoking status is set to zero with probability 0.16 based on the reported efficacy of a behavioral smoking cessation intervention described by Hoffman et al. and tested in the WIHS.[121]

Portfolio 4B: Under this portfolio, both Portfolio 2B and 3B are implemented.

Identification Conditions

To consistently estimate our parameters of interest with the parametric g-computation algorithm, we assume a set of sufficient conditions. First, we assume no measurement error of exposures, covariates, or outcome.[107] More specifically, we assume that the exposures and outcome are measured without error while covariates might be mismeasured but that the observed values are those that inform the intervention plan. For example, a laboratory measurement of CD4 cell count might be 220 for a patient whose true CD4 cell count is 235; however, the 220 value the clinician sees is the value that is used to determine

whether or not to initiate ART. Second, we assume counterfactual consistency (also described as treatment version irrelevance and no interference).[108,144] If the observed exposure of woman i is equal to the value of the exposure indicated by the intervention of interest (i.e., $A_i = a^g$), then the observed outcome of woman i is equal to the outcome that would have been observed if exposure was set to the value indicated by the intervention plan. Third, we assume conditional exchangeability or no unmeasured confounding;[135] the probability of exposures conditional on the set of modeled covariates is independent of the potential outcome. Fourth, we assume positivity, also described as the experimental treatment assignment assumption;[111] the conditional probability of receiving exposure is nonzero within all strata of covariates. Finally, we assume correct specification of parametric models used to estimate the g-computation algorithm formula.

Estimation Process

We followed the previously outlined steps for parametrically estimating the g-computation formula.[68,69,71] Briefly, we fit pooled person-period logistic and linear models to the observed data across all visits $j \geq 1$ to estimate the conditional probability or density of each exposure, time-varying covariate, and outcome (Step 1). We then drew a large Monte Carlo sample of $N = 100{,}000$ women at baseline (i.e., $j = 0$) with replacement (Step 2). In the Monte Carlo sample, we used the estimated conditional probabilities or densities and observed values of exposures and covariates at the baseline visit to simulate the values of exposures, covariates, and outcomes at $j = 1$ by drawing values from a Bernoulli distribution in the case of binary variables and multinomial distributions in the case of categorical variables. We then used the estimated conditional probabilities or densities and simulated values of exposures, covariates, and outcomes at $j = 1$ to simulate the values of these variables at $j = 2$, continuing this process for a maximum of 16 follow-up visits (Step 3). After completing the Monte Carlo simulation, we estimated the cumulative incidence of mortality at each time point using the complement of the Kaplan-Meier[125] estimator of the survival function (Step 4). If parametric models are correctly specified, the cumulative incidence of mortality estimated in the Monte Carlo dataset under the natural course (i.e.,

no intervention on exposures) should be equivalent to the cumulative incidence of mortality in the WIHS dataset. Similarly, the distribution of each exposure and time-varying covariate in the Monte Carlo sample should be equivalent to the distributions in the observed WIHS data.

However, an important point is that we did not include a censoring model in our parametric g-computation algorithm. Rather, we generated follow-up data for women at all visits at which they remained alive and had completed less than 8 years under study since analysis baseline. Thus, in our generation of data under the "natural course," we use the definition of "natural course" provided by Young et al.[78] in which loss to follow-up is eliminated but there is no intervention on exposure. Thus, we assume conditional exchangeability for those lost and not lost to follow-up based on modeled covariate history. Eliminating loss to follow-up in the g-computation simulated data may therefore result in a different number of observed events than were observed in the WIHS data. These two numbers should not be expected to be equal. However, the cumulative incidence function may still be equivalent, assuming no informative censoring.

To achieve good model fit, we repeated the above steps using flexible specifications of continuous and categorical variables, interaction terms, and various sets of covariates for each model until we 1) had included all hypothesized confounders relevant to each exposure, and 2) were able to replicate the cumulative incidence observed in the WIHS sample in the Monte Carlo sample. The variables and models that were selected for final analyses are described below.

Time-fixed covariates measured at baseline ($\boldsymbol{W}_{j=0}$) included age, race, level of education, history of injection drug use (IDU), prior exposure to dual therapy zidovudine (AZT), wave of WIHS enrollment, CD4 count, viral load, and smoking status. In final models, age was included as linear and squared terms (age_0, age_0^2). Race was included as a binary indicator of Black race. Level of education was included as indicators for having completed high school ($Education_{0,1}$) or having some degree of college education ($Education_{0,2}$) with the reference category defined as not having completed high school. History of injection drug use, prior exposure to AZT, and smoking were included as binary indicators. Wave of

WIHS enrollment was included as a binary indicator of having been enrolled in WIHS during the first recruitment wave between 1994 – 1995 ($FirstWave_0$). CD4 and viral load at baseline were included as categorical indicators using the same definitions described by Westreich et al.[68] where CD4 categories were <200 cells/ml ($CD4_{0,1}$) and >350 cells/ml ($CD4_{0,2}$) with a reference category of 200-350 cells/ml and viral load categories were ≤400 copies/ml ($VL_{0,1}$) and >10,000 copies/ml ($VL_{0,2}$) with a reference category of 401-10,000 copies/ml.

Time-varying covariates and exposures were CD4 cell count, viral load, depression, antiretroviral therapy, alcohol consumption, and smoking. CD4 cell count was included as a linear term, viral load was included as a binary indicator of having a detectable viral load (DVL), and depression was included as a binary indicator of having a depressive symptom score indicating depression according to CES-D scale scoring criteria.[124] Antiretroviral therapy was included as binary indicators of being on a combination ART treatment regimen (ART), having promptly initiated ART within 6 months of baseline ($PromptART$), and having initiated ART in the modern treatment period on/after 1 October 2001 ($ModernART$). ART exposure was also included as indicator terms for time on ART at each visit to allow for interactions between ART and time ($ARTvisits$). Alcohol consumption was included as a binary indicator of any drinking ($AnyD$) as well as indicators for the average number of alcohol consumed per week ($Drinklevel$) where 1=1 drink per week, 2=2 drinks per week, 3=3 drinks per week, 4=4-7 drinks per week, 5=8-12 drinks per week, and 6=over 12 drinks per week. Smoking was included as a binary indicator of smoking since the last visit.

We also assessed interactions for smoking and high alcohol consumption as well as hepatitis C virus and high alcohol consumption. High alcohol consumption was defined as a binary indicator of consuming over 3 drinks per week. Hepatitis C virus was defined as having a positive antibody test and positive or missing RNA test at baseline. Ultimately, we excluded the interaction term between smoking and high alcohol consumption in our final models based on assessment of model fit. Estimates of the cumulative incidence functions and risks obtained using models with and without the interaction term

were similar but models were less stable with the inclusion of the interaction term and a likelihood ratio test did not support significant interaction (p-value 0.7). In each model, we included time (i.e., visit number) as restricted quadratic spline terms with 4 knots (k, k^1, k^2, k^3).

Models: Final models used to estimate the conditional probability or density of each variable are shown below where $expit(\alpha) = \frac{\exp(\alpha)}{[1+\exp(\alpha)]}$ and error terms ε are normal.

Death

$$
\begin{aligned}
&P\left(Y_{k+1}=1 \middle| \bar{A}_k=\bar{a}_k, \bar{D}_k=\bar{d}_k,\ \bar{S}_k=\bar{s}_k,\ \overline{\boldsymbol{Z}}_k=\overline{\boldsymbol{z}}_k,\ \overline{\boldsymbol{W}}_k=\overline{\boldsymbol{w}}_k, \bar{C}_k=\bar{Y}_k=0\right) \\
&\quad = expit\big[\alpha_0+\alpha_1 k+\alpha_2 k^1+\alpha_3 k^2+\alpha_4 k^3+\alpha_5 age_0+\alpha_6 age_0^2+\alpha_7 Race_0+\alpha_8 IDU_0+\ \alpha_9 Education_{0,1} \\
&\quad +\ \alpha_{10} Education_{0,2}+\ \alpha_{11} AZT_0+\ \alpha_{12} CD4_{0,1}+\alpha_{13} CD4_{0,2}+\alpha_{14} VL_{0,1}+\alpha_{15} VL_{0,2}+\alpha_{16} Smoke_0 \\
&\quad +\alpha_{17} CD4_k+\alpha_{18} CD4_{k-1}+\alpha_{19} DVL_k+\alpha_{20} DVL_{k-1}+\alpha_{21} ART_k+\alpha_{22} ART_{k-1}+\alpha_{23} ModernART_k \\
&\quad +\alpha_{24} PromptART_k+\alpha_{25} I(ARTvisits_k=2) \\
&\quad +\alpha_{26} I(ARTvisits_k=3)+\alpha_{27} I(ARTvisits_k=4)+\alpha_{28} I(ARTvisits_k=5)+\alpha_{29} I(ARTvisits_k=6) \\
&\quad +\alpha_{30} I(ARTvisits_k=7)+\alpha_{31} I(ARTvisits_k=8)+\alpha_{32} I(ARTvisits_k=9) \\
&\quad +\alpha_{33} I(ARTvisits_k=10)+\alpha_{34} I(ARTvisits_k=11)+\alpha_{35} I(ARTvisits_k=12) \\
&\quad +\alpha_{36} I(ARTvisits_k=13) \\
&\quad +\alpha_{37} I(ARTvisits_k=14)+\alpha_{38} I(ARTvisits_k=15)+\ \alpha_{39} I(ARTvisits_k=16)+\alpha_{40} AnyD_k \\
&\quad +\ \alpha_{41} Drinklevel_k+\alpha_{42} High\,Alcohol_{k-1}+\alpha_{43} Depressed_k+\alpha_{44}(ART*Depressed)_k+\alpha_{45} Smoke_k \\
&\quad +\alpha_{46} Smoke_{k-1}+\ \alpha_{47}(HCV*High\,Alcohol)_k\big]
\end{aligned}
$$

Alcohol consumption

Two models were specified for alcohol consumption. First, we estimated the probability of any alcohol consumption as:

$$
\begin{aligned}
&P\left(AnyD_k=1 \middle|\ \bar{D}_{k-1}=\bar{d}_k,\ \bar{S}_k=\bar{s}_k,\ \overline{\boldsymbol{Z}}_{k-1}=\overline{\boldsymbol{z}}_{k-1},\ \overline{\boldsymbol{W}}_k=\overline{\boldsymbol{w}}_k, \bar{C}_k=\bar{Y}_k=0\right) \\
&\quad = expit(\beta_0+\beta_1 k+\beta_2 k^1+\beta_3 k^2+\beta_4 k^3+\beta_5 age_0+\beta_6 age_0^2+\beta_7 Race_0+\beta_8 IDU_0+\ \beta_9 Depressed_{k-1} \\
&\quad +\beta_{10} Drinklevel_{k-1})
\end{aligned}
$$

Among those who consumed alcohol, we estimated the probability of each drinking level from a generalized logit model where probabilities of each drinking level are given by $\frac{I(Level=level)\times exp\ (\alpha)}{1+\exp\ (\alpha)}$; the numerator exp function includes all beta coefficients for the drinking level being estimated while the denominator exp function includes beta coefficients for all non-reference categories. In our case, the

reference category was a drinking level of 1. Here, we use a drinking level of 6 as an example of our model.

$$
\begin{aligned}
P\big(D_k = 6 \big| \bar{D}_{k-1} = \bar{d}_{k,}\ \bar{S}_k = \bar{s}_{k,}\ \overline{\boldsymbol{Z}}_{k-1} = \bar{\boldsymbol{z}}_{k-1},\ \overline{\boldsymbol{W}}_k = \bar{\boldsymbol{w}}_k, \bar{C}_k = \bar{Y}_k = 0\big) \\
&= I(Level = 6) \\
&\times \{exp\big(\beta_{0,6} + \beta_{1,6}k + \beta_{2,6}k^1 + \beta_{3,6}k^2 + \beta_{4,6}k^3 + \beta_{5,6}age_0 + \beta_{6,6}age_0^2 + \beta_{7,6}Race_0 + \beta_{8,6}IDU_0 \\
&+ \beta_{9,6}Depressed_{k-1} + \beta_{10,6}Drinklevel_{k-1}\big) \\
&/\big(1 \\
&+ exp\big[\beta_{0,6} + \beta_{0,5} + \beta_{0,4} + \beta_{0,3} + \beta_{0,2} + \beta_{1,6}k + \beta_{1,5}k + \beta_{1,4}k + \beta_{1,3}k + \beta_{1,2}k + \beta_{2,6}k^1 \\
&+ \beta_{2,5}k^1 + \beta_{2,4}k^1 + \beta_{2,3}k^1 + \beta_{2,2}k^1 + \beta_{2,1}k^1 + \beta_{3,6}k^2 \\
&+ \beta_{3,5}k^2 + \beta_{3,4}k^2 + \beta_{3,3}k^2 + \beta_{3,2}k^2 + \beta_{4,6}k^3 + \beta_{4,5}k^3 + \beta_{4,4}k^3 + \beta_{4,3}k^3 + \beta_{4,2}k^3 + \beta_{5,6}age_0 \\
&+ \beta_{5,5}age_0 + \beta_{5,4}age_0 + \beta_{5,3}age_0 + \beta_{5,2}age_0 + \beta_{6,6}age_0^2 + \beta_{6,5}age_0^2 + \beta_{6,4}age_0^2 + \beta_{6,3}age_0^2 \\
&+ \beta_{6,2}age_0^2 + \beta_{7,6}Race_0 + \beta_{7,5}Race_0 + \beta_{7,4}Race_0 + \beta_{7,3}Race_0 + \beta_{7,2}Race_0 + \beta_{8,6}IDU_0 + \beta_{8,5}IDU_0 \\
&+ \beta_{8,4}IDU_0 + \beta_{8,3}IDU_0 + \beta_{8,2}IDU_0 + \beta_{9,6}Depressed_{k-1} + \beta_{9,5}Depressed_{k-1} + \beta_{9,4}Depressed_{k-1} \\
&+ \beta_{9,3}Depressed_{k-1} + \beta_{9,2}Depressed_{k-1} + \beta_{10,6}Drinklevel_{k-1} + \beta_{10,5}Drinklevel_{k-1} \\
&+ \beta_{10,4}Drinklevel_{k-1} + \beta_{10,3}Drinklevel_{k-1} + \beta_{10,2}Drinklevel_{k-1}\big]\big)\}
\end{aligned}
$$

The probability of each drinking level except for that of the reference category was calculated in the same manner using the appropriate beta coefficients in the numerator. The probability of the reference drinking level was calculated as 1 minus the sum of all other drinking level probabilities.

<u>Smoking (stratified by smoking at $k-1$):</u>

$$
\begin{aligned}
P\big(S_k = 1 \big|\ \bar{D}_{k-1} = \bar{d}_{k,}\ \bar{S}_k = \bar{s}_{k,}\ \overline{\boldsymbol{Z}}_{k-1} = \bar{\boldsymbol{z}}_{k-1},\ \overline{\boldsymbol{W}}_k = \bar{\boldsymbol{w}}_k, \bar{C}_k = \bar{Y}_k = 0\big) \\
&= expit\big[\gamma_0 + \gamma_1 k + \gamma_2 k^1 + \gamma_3 k^2 + \gamma_4 k^3 + \gamma_5 age_0 + \gamma_6 age_0^2 + \gamma_7 Race_0 + \gamma_8 IDU_0 + \gamma_9 Education_{0,1} \\
&+ \gamma_{10} Education_{0,2} + \gamma_{11} I(Drinklevel_{k-1} = 0) + \gamma_{12} I(Drinklevel_{k-1} = 1) + \gamma_{13} I(Drinklevel_{k-1} = 2) \\
&+ \gamma_{14} I(Drinklevel_{k-1} = 3) + \gamma_{15} I(Drinklevel_{k-1} = 4) + \gamma_{16} I(Drinklevel_{k-1} = 5) \\
&+ \gamma_{17} I(Drinklevel_{k-1} = 6)\big]
\end{aligned}
$$

<u>Depression:</u>

$$
\begin{aligned}
P\big(Depressed_k = 1 \big| \bar{A}_k = \bar{a}_k\ , \bar{D}_k = \bar{d}_{k,}\ \bar{S}_k = \bar{s}_{k,}\ \overline{\boldsymbol{Z}}_{k-1} = \bar{\boldsymbol{z}}_{k-1},\ \overline{\boldsymbol{W}}_k = \bar{\boldsymbol{w}}_k, \bar{C}_k = \bar{Y}_k = 0\big) \\
&= expit\big(\theta_0 + \theta_1 k + \theta_2 k^1 + \theta_3 k^2 + \theta_4 k^3 + \theta_5 age_0 + \theta_6 age_0^2 + \theta_7 Race_0 + \theta_8 CD4_{0,1} + \theta_9 CD4_{0,2} \\
&+ \theta_{10} ART_k + \theta_{11} ART_{k-1} + \theta_{12} Depressed_{k-1} + \theta_{13} Depressed_{k-2} + \theta_{14} Smoke_{k-1} \\
&+ \theta_{15} High\ Alcohol_{k-1}\big)
\end{aligned}
$$

Detectable viral load (stratified by detectable viral load at $k-1$):

$$P\left(DVL_k = 1 \middle| \bar{A}_k = \bar{a}_k, \bar{D}_k = \bar{d}_k, \ \bar{S}_k = \bar{s}_k, \ \bar{\boldsymbol{Z}}_{k-1} = \bar{\boldsymbol{z}}_{k-1}, \ \bar{\boldsymbol{W}}_k = \bar{\boldsymbol{w}}_k, \bar{C}_k = \bar{Y}_k = 0\right)$$
$$= expit\big[\theta_0 + \theta_1 k + \theta_2 k^1 + \theta_3 k^2 + \theta_4 k^3 + \theta_5 age_0 + \theta_6 age_0^2 + \theta_7 Race_0 + \theta_8 CD4_{0,1} + \ \theta_9 CD4_{0,2}$$
$$+ \ \theta_{10} VL_{0,1} + \theta_{11} VL_{0,2} \ + \theta_{12} FirstWave_0 \ + \theta_{13} ART_k \ + \theta_{14} ModernART_k + \theta_{15} CD4_{k-1}$$
$$+ \ \theta_{16} I(Drinklevel_k = 0) + \ \theta_{17} I(Drinklevel_k = 1) + \theta_{18} I(Drinklevel_k = 2)$$
$$+ \ \theta_{19} I(Drinklevel_k = 3) \ + \ \theta_{20} I(Drinklevel_{k-1} = 4) \ + \ \theta_{21} I(Drinklevel_{k-1} = 5)$$
$$+ \ \theta_{22} I(Drinklevel_{k-1} = 6)\big]$$

CD4:

$$f\left(\ln\left(CD4_k\right) \middle| \bar{A}_k = \bar{a}_k, \bar{D}_k = \bar{d}_k, \ \bar{S}_k = \bar{s}_k, \bar{\boldsymbol{Z}}_{k-1} = \bar{\boldsymbol{z}}_{k-1}, \ \bar{\boldsymbol{W}}_k = \bar{\boldsymbol{w}}_k, \bar{C}_k = \bar{Y}_k = 0\right) =$$
$$= exp\big[\rho_0 + \rho_1 k + \rho_2 k^1 + \rho_3 k^2 + \rho_4 k^3 + \rho_5 age_0 + \rho_6 age_0^2 + \rho_7 Race_0 + \rho_8 CD4_{0,1} + \ \rho_9 CD4_{0,2}$$
$$+ \ \rho_{10} VL_{0,1} + \rho_{11} VL_{0,2} \ + \rho_{13} ART_k \ + \rho_{14} ART_{k-1} + \rho_{15} ModernART_k$$
$$+ \rho_{16} PromptART_k + \rho_{17} lnCD4_{k-1} + \ \rho_{18} lnCD4_{k-2} + \rho_{19} DVL_{k-1} + \ \rho_{20} DVL_{k-2}$$
$$+ \ \rho_{21} I(Drinklevel_k = 0) + \ \rho_{22} I(Drinklevel_k = 1) + \rho_{23} I(Drinklevel_k = 2)$$
$$+ \ \rho_{24} I(Drinklevel_k = 3) \ + \ \rho_{25} I(Drinklevel_{k-1} = 4) \ + \ \rho_{26} I(Drinklevel_{k-1} = 5)$$
$$+ \ \rho_{27} I(Drinklevel_{k-1} = 6)\big]$$

To estimate the cumulative incidence of mortality in the WIHS under each intervention portfolio, we repeated the steps described in the Estimation Process for each intervention. However, we altered the value of the exposures in the Monte Carlo sample in accordance with the criteria set forth by the intervention definitions. We estimated the cumulative incidence of mortality with the complement of the Kaplan-Meier[125] survival function in each dataset. To calculate risk differences and risk ratios contrasting the effect of each intervention portfolio with the base scenario of intervening only on prompt initiation of modern ART, we concatenated datasets to subtract mortality risks under each intervention from the risks estimated in the base scenario.

To obtain 95% confidence intervals, we repeated all steps for each intervention portfolio in 1000 nonparametric bootstrap resamples[126] taken of the original WIHS data. The standard error was calculated based on the standard deviation of estimated risks and effect measures across all bootstrap resamples.

APPENDIX B: SUPPLEMENTAL TEXT

Data generation for motivating example

Let A represent exposure to long (i.e., $A = 1$) versus short (i.e., $A = 1$) duration of opioid prescription, W represent history of substance use (history: $W = 1$; no history: $W = 0$), Z represent lack of engagement in regular HIV care (lack of engagement: $Z = 1$; engagement: $Z = 0$), $R = 1$ represent observed exposure information (i.e., selection into the sample of complete cases), $Y^{a=0}$ represent the potential outcome for emergency room visit or death under short term opioid exposure, $Y^{a=1}$ represent the potential outcome for emergency room visit or death under long term opioid exposure, and Y represent the factual outcome for emergency room visit or death given the exposure received.

We generated 10,000 simulation trials of n = 1 to 1,623 independent and identically distributed sets $\{W, A, Z, R, Y^{a=0}, Y^{a=1}, Y\}_n$ representing the N = 1,623 members in the full population of interest. For each record, covariate values were simulated in the following order:

W drawn from a Bernoulli distribution with a marginal prevalence of 0.44

Z drawn from a Bernoulli distribution with a marginal prevalence of 0.4

R drawn from a Bernoulli distribution as a function of Z and with a marginal prevalence of 0.3

$Y^{a=0}$ drawn from a Bernoulli distribution as a function of $W$and Z and with a marginal incidence of 0.19

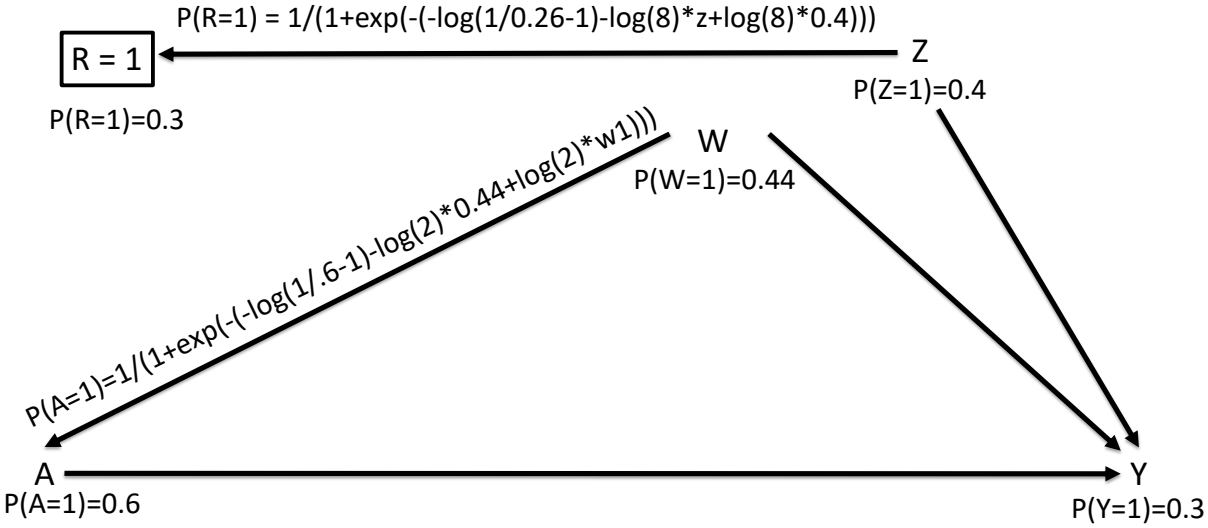
$Y^{a=1}$ drawn from a Bernoulli distribution as a function of $W$and Z and with a marginal incidence of 0.37

A drawn from a Bernoulli distribution as a function of W and with a marginal prevalence of 0.6

Y set according to realized value a

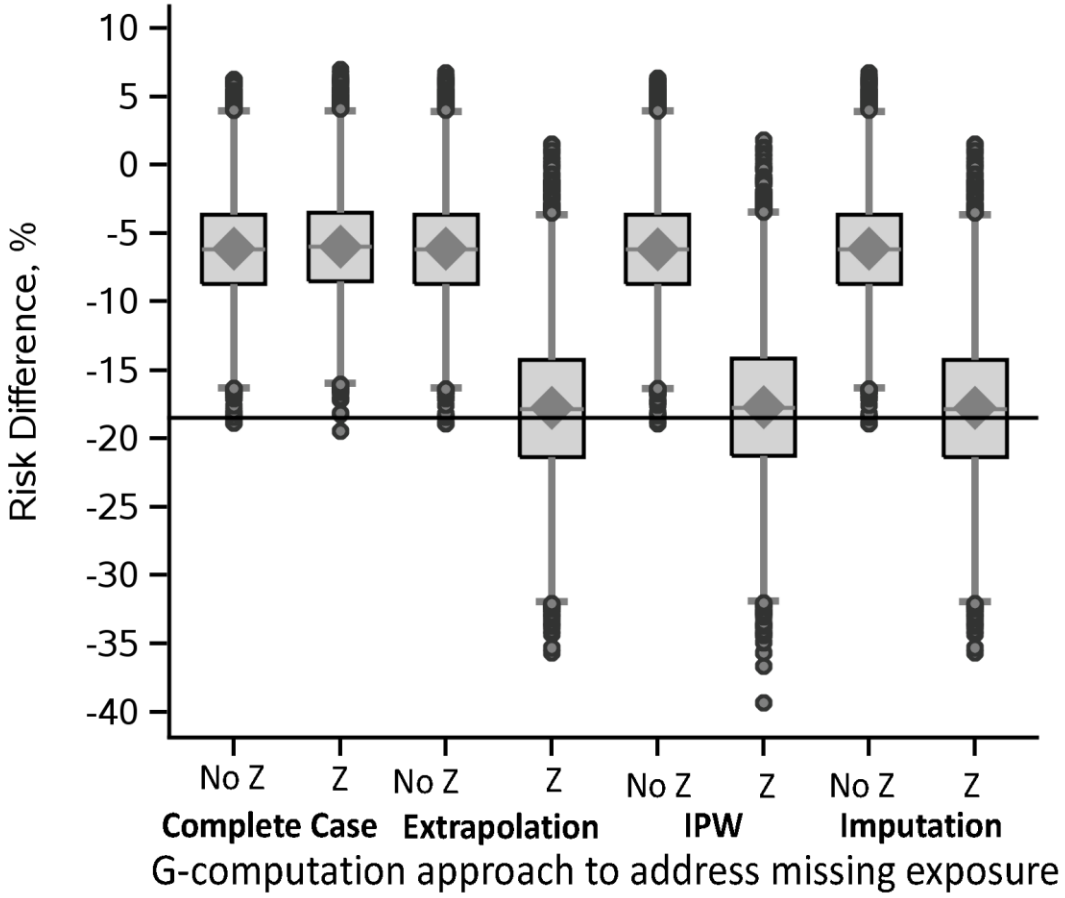
SAS code used to generate the cohort and a graphical representation of the data generating mechanism with chosen parameters are shown below.

```
data cohort;
        call streaminit(123);
        do trial = 1 to 10000;
        do n =  1 to 1623 ;
        w = rand("BERNOULLI",0.44);
        *History of substance use;
        z = rand("BERNOULLI",0.4);
        *Not being established in regular HIV care;
        r = rand("Bernoulli",1/(1+exp(-(-log(1/0.26-1)-log(8)*z+log(8)*0.4))));
        *Complete data;
        y0 = rand("BERNOULLI",1/(1+exp(-(-log(1/.19-1)-log(1.5)*0.44+log(1.5)*w1-
        log(1.5)*0.4+log(1.5)*z))));
        *Potential outcome for ER visit or death under short duration opioid prescription;
        y1 = rand("BERNOULLI",1/(1+exp(-(-log(1/.35-1)-log(1.5)*0.44+log(1.5)*w1-
        log(10)*0.4+log(10)*z))));
        *Potential outcome for ER visit or death under long duration opioid prescription;
        a = rand("BERNOULLI",1/(1+exp(-(-log(1/.6-1)-log(2)*0.44+log(2)*w1))));
        *Long duration prescription;
        if a then y=y1; else y=y0;
        *Factual outcome, i.e., observed potential outcome given actual exposure;
        output; end; end; run;
```

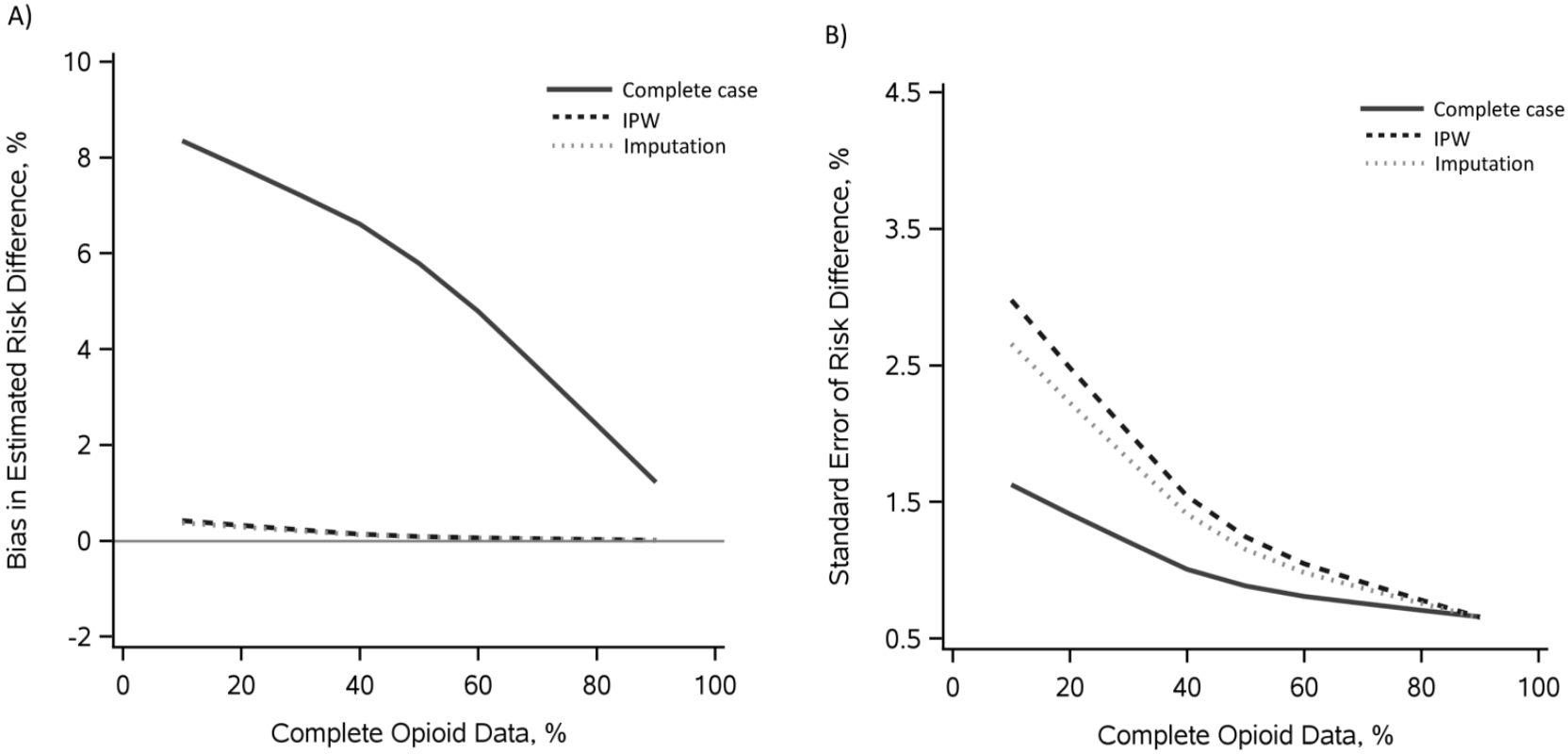


$P(Y^0)=1/(1+\exp(-(-\log(1/.19-1)-\log(1.5)*0.44+\log(1.5)*w1-\log(1.5)*0.4+\log(1.5)*z)))$
$P(Y^1)=1/(1+\exp(-(-\log(1/.35-1)-\log(1.5)*0.44+\log(1.5)*w1-\log(10)*0.4+\log(10)*z)))$

Appendix Figure 1. Risk difference for the average treatment effect estimated by g-computation approaches including and excluding covariate Z.



Appendix Figure 2. Bias (Panel A) and standard error (Panel B) for the effect of shortening opioid prescription duration among 7,500 HIV-positive women estimated by three g-computation approaches to address missing opioid data, 10,000 simulation experiments with complete opioid data varying from 10-90%.



REFERENCES

1. Crum NF, Riffenburgh RH, Wegner S, et al. Comparisons of causes of death and mortality rates among HIV-infected persons: analysis of the pre-, early, and late HAART (highly active antiretroviral therapy) eras. *J Acquir Immune Defic Syndr*. 2006;41(2):194-200. doi:00126334-200602010-00011 [pii].

2. Palella FJ, Baker RK, Moorman AC, et al. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. *J Acquir Immune Defic Syndr*. 2006;43(1):27-34. doi:10.1097/01.qai.0000233310.90484.16.

3. French AL, Gawel SH, Hershow R, et al. Trends in mortality and causes of death among women with HIV in the United States: a 10-year study. *J Acquir Immune Defic Syndr*. 2009;51(4):399-406. doi:10.1097/QAI.0b013e3181acb4e5.

4. Schwarcz SK, Vu A, Hsu LC, Hessol NA. Changes in Causes of Death Among Persons with AIDS: San Francisco, California, 1996-2011. *AIDS Patient Care STDS*. 2014;28(10):517-523. http://www.ncbi.nlm.nih.gov/pubmed/25275657.

5. Buchacz K, Baker RK, Palella FJ, et al. Disparities in prevalence of key chronic diseases by gender and race/ethnicity among antiretroviraltreated HIV-infected adults in the US. *Antivir Ther*. 2013;18(1):65-75. doi:10.3851/IMP2450.

6. Hemkens LG, Bucher HC. HIV infection and cardiovascular disease. *Eur Heart J*. 2014;35(21):1373-1381. doi:10.1093/eurheartj/eht528.

7. Hsue PY, Lo JC, Franklin A, et al. Progression of Atherosclerosis as Assessed by Carotid Intima-Media Thickness in Patients With HIV Infection. *Circulation*. 2004;109(13):1603-1608. doi:10.1161/01.CIR.0000124480.32233.8A.

8. Lo J, Abbara S, Shturman L, et al. Increased prevalence of subclinical coronary atherosclerosis detected by coronary computed tomography angiography in HIV-infected men. *AIDS*. 2010;24(2):243-253. doi:10.1097/QAD.0b013e328333ea9e.

9. Freiberg MS, Chang C-CH, Kuller LH, et al. HIV Infection and the Risk of Acute Myocardial Infarction. *JAMA Intern Med*. 2013;173(8):614. doi:10.1001/jamainternmed.2013.3728.

10. Rodriguez-Penney AT, Iudicello JE, Riggs PK, et al. Co-morbidities in persons infected with HIV: increased burden with older age and negative effects on health-related quality of life. *AIDS Patient Care STDS*. 2013;27(1):5-16. doi:10.1089/apc.2012.0329.

11. Price JC, Thio CL. Liver disease in the HIV-infected individual. *Clin Gastroenterol Hepatol*. 2010;8(12):1002-1012. doi:10.1016/j.cgh.2010.08.024.

12. Joshi D, O'Grady J, Dieterich D, Gazzard B, Agarwal K. Increasing burden of liver disease in patients with HIV infection. *Lancet*. 2011;377(9772):1198-1209. doi:10.1016/S0140-6736(10)62001-6.

13. Crothers K, Butt AA, Gibert CL, Rodriguez-Barradas MC, Crystal S, Justice AC. Increased COPD among HIV-positive compared to HIV-negative veterans. *Chest*. 2006;130(5):1326-1333. doi:10.1378/chest.130.5.1326.

14. Crothers K, Thompson BW, Burkhardt K, et al. HIV-Associated Lung Infections and Complications in the Era of Combination Antiretroviral Therapy. *Proc Am Thorac Soc*. 2011;8(3):275-281. doi:10.1513/pats.201009-059WR.

15. Raynaud C, Roche N, Chouaid C. Interactions between HIV infection and chronic obstructive pulmonary disease: Clinical and epidemiological aspects. *Respir Res*. 2011;12(1):117. doi:10.1186/1465-9921-12-117.

16. Kim DJ, Westfall AO, Chamot E, et al. Multimorbidity patterns in HIV-infected patients: the role of obesity in chronic disease clustering. *J Acquir Immune Defic Syndr*. 2012;61(5):600-605. doi:10.1097/QAI.0b013e31827303d5.

17. Kooij KW, Wit FWNM, Schouten J, et al. HIV infection is independently associated with frailty in middle-aged HIV type 1-infected individuals compared with similar but uninfected controls. *AIDS*. 2016;30(2):241-250. doi:10.1097/QAD.0000000000000910.

18. Guaraldi G, Orlando G, Zona S, et al. Premature age-related comorbidities among HIV-infected persons compared with the general population. *Clin Infect Dis*. 2011;53(11):1120-1126. doi:10.1093/cid/cir627.

19. Pathai S, Bajillan H, Landay AL, High KP. Is HIV a model of accelerated or accentuated aging? *Journals Gerontol - Ser A Biol Sci Med Sci*. 2014;69(7):833-842. doi:10.1093/gerona/glt168.

20. Samji H, Cescon A, Hogg RS, et al. Closing the Gap: Increases in Life Expectancy among Treated HIV-Positive Individuals in the United States and Canada. Okulicz JF, ed. *PLoS One*. 2013;8(12):e81355. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0081355. Accessed October 6, 2015.

21. Wada N, Jacobson LP, Cohen M, French A, Phair J, Muñoz A. Cause-specific life expectancies after 35 years of age for human immunodeficiency syndrome-infected and human immunodeficiency syndrome-negative individuals followed simultaneously in long-term cohort studies, 1984-2008. *Am J Epidemiol*. 2013;177(2):116-125. doi:10.1093/aje/kws321.

22. Harrison KM, Song R, Zhang X. Life expectancy after HIV diagnosis based on national HIV surveillance data from 25 states, United States. *J Acquir Immune Defic Syndr*. 2010;53(1):124-130. http://www.ncbi.nlm.nih.gov/pubmed/19730109. Accessed February 22, 2016.

23. Mdodo R, Frazier EL, Dube SR, et al. Cigarette smoking prevalence among adults with HIV compared with the general adult population in the United States: Cross-sectional surveys. *Ann Intern Med*. 2015;162(5):335-344. doi:10.7326/M14-0954.

24. Galvan FH, Bing EG, Fleishman J a, et al. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: results from the HIV Cost and Services Utilization Study. *J Stud Alcohol*. 2002;63(2):179-186. doi:10.15288/jsa.2002.63.179.

25. Chander G, Josephs J, Fleishman JA, et al. Alcohol use among HIV-infected persons in care: Results of a multi-site survey. *HIV Med*. 2008;9(4):196-202. doi:10.1111/j.1468-1293.2008.00545.x.

26. Samet JH, Cheng DM, Libman H, Nunes DP, Alperen JK, Saitz R. Alcohol consumption and HIV disease progression. *J Acquir Immune Defic Syndr*. 2007;46(2):194-199. doi:10.1097/QAI.0b013e318142aabb.

27. Neblett RC, Hutton HE, Lau B, McCaul ME, Moore RD, Chander G. Alcohol consumption among HIV-infected women: impact on time to antiretroviral therapy and survival. *J Womens Health (Larchmt)*. 2011;20(2):279-286. doi:10.1089/jwh.2010.2043.

28. Phillips AN, Neaton J, Lundgren JD. The role of HIV in serious diseases other than AIDS. *AIDS*. 2008;22(18):2409-2418. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2679976&tool=pmcentrez&rendertype=abstract. Accessed November 15, 2015.

29. Kuller LH, Tracy R, Belloso W, et al. Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. *PLoS Med*. 2008;5(10):1496-1508. doi:10.1371/journal.pmed.0050203.

30. Warriner AH, Burkholder GA, Overton ET. HIV-related metabolic comorbidities in the current ART era. *Infect Dis Clin North Am*. 2014;28(3):457-476. doi:10.1016/j.idc.2014.05.003.

31. Mukamal KJ. The effects of smoking and drinking on cardiovascular disease and risk factors. *Alcohol Res Heal*. 2006.

32. Bellentani S, Saccoccio G, Costa G, et al. Drinking habits as cofactors of risk for alcohol induced liver damage. *Gut*. 1997. doi:10.1136/gut.41.6.845.

33. Hamabe A, Uto H, Imamura Y, et al. Impact of cigarette smoking on onset of nonalcoholic fatty liver disease over a 10-year period. *J Gastroenterol*. 2011. doi:10.1007/s00535-011-0376-z.

34. Lee J, Taneja V, Vassallo R. Cigarette smoking and inflammation: Cellular and molecular mechanisms. *J Dent Res*. 2012. doi:10.1177/0022034511421200.

35. McEvoy JW, Nasir K, Defilippis AP, et al. Relationship of cigarette smoking with inflammation and subclinical vascular disease: The multi-ethnic study of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2015. doi:10.1161/ATVBAHA.114.304960.

36. Imhof A, Froehlich M, Brenner H, Boeing H, Pepys MB, Koenig W. Effect of alcohol consumption on systemic markers of inflammation. *Lancet*. 2001. doi:10.1016/S0140-6736(00)04170-2.

37. Stanhope B-J, Burdette WJ, Cochran WG, et al. *Report of the Advisory Committee to the Surgeon General of the Public Health Service*.; 1964.

38. U.S. Department of Health and Human Services. *A Report of the Surgeon General: How Tobacco Smoke Causes Disease: What It Means to You. U.S.*; 2010. doi:Dec 1 2014.

39. Eliasson B. Cigarette smoking and diabetes. *Prog Cardiovasc Dis*. 2003. doi:10.1053/pcad.2003.00103.

40. Maddatu J, Anderson-Baucum E, Evans-Molina C. Smoking and the risk of type 2 diabetes. *Transl Res*. 2017. doi:10.1016/j.trsl.2017.02.004.

41. Rahmanian S, Wewers ME, Koletar S, Reynolds N, Ferketich A, Diaz P. Cigarette Smoking in the HIV-Infected Population. *Proc Am Thorac Soc.* 2011;8(3):313-319. doi:10.1513/pats.201009-058WR.

42. Helleberg M, Afzal S, Kronborg G, et al. Mortality attributable to smoking among HIV-1-infected individuals: A nationwide, population-based cohort study. *Clin Infect Dis.* 2013;56(5):727-734. doi:10.1093/cid/cis933.

43. Calvo-Sánchez M, Perelló R, Pérez I, et al. Differences between HIV-infected and uninfected adults in the contributions of smoking, diabetes and hypertension to acute coronary syndrome: Two parallel case-control studies. *HIV Med*. 2013;14(1):40-48. doi:10.1111/j.1468-1293.2012.01057.x.

44. Gingo MR, Morris A. Pathogenesis of HIV and the Lung. *Curr HIV/AIDS Rep*. 2012:42-50. doi:10.1007/s11904-012-0140-x.

45. Klitzman RL, Kirshenbaum SB, Dodge B, et al. Intricacies and inter-relationships between HIV disclosure and HAART: a qualitative study. *AIDS Care*. 2004;16:628-640. doi:10.1080/09540120410001716423.

46. Westreich D, Cates J, Cohen M, et al. Smoking, HIV, and risk of pregnancy loss. *AIDS*. 2017;31(4):553-560. doi:10.1097/QAD.0000000000001342.

47. Justice AC, McGinnis KA, Tate JP, et al. Risk of mortality and physiologic injury evident with lower alcohol exposure among HIV infected compared with uninfected men. *Drug Alcohol Depend*. 2016;161:95-103. doi:10.1016/j.drugalcdep.2016.01.017.

48. Tien PC, Choi AI, Zolopa AR, et al. Inflammation and mortality in HIV-infected adults: analysis of the FRAM study cohort. *J Acquir Immune Defic Syndr*. 2010;55(3):316-322. doi:10.1097/QAI.0b013e3181e66216.

49. Westreich D. From exposures to population interventions: Pregnancy and response to HIV therapy. *Am J Epidemiol*. 2014;179(7):797-806. doi:10.1093/aje/kwt328.

50. Westreich D, Edwards JK, Rogawski ET, Hudgens MG, Stuart EA, Cole SR. Causal impact: Epidemiological approaches for a public health of consequence. *Am J Public Health*. 2016. doi:10.2105/AJPH.2016.303226.

51. Westreich D. From Patients to Policy: Population Intervention Effects in Epidemiology. *Epidemiology*. 2017;28(4). doi:10.1097/EDE.0000000000000648.

52. Westreich D, Mooney SJ. Number (of Whom?) Needed to Treat (with What?): Exposures, Population Interventions, and the Number Needed to Treat. *Epidemiology*. 2019. doi:10.1097/EDE.0000000000001061.

53. Cahill K, Stevens S, Perera R, Lancaster T. Pharmacological interventions for smoking cessation: an overview and network meta-analysis. *Cochrane Database Syst Rev*. 2013;5(5):CD009329. doi:10.1002/14651858.CD009329.pub2.

54. Yeung E, Long S, Bachi B, Lee J, Chao Y. The psychiatric effects of varenicline on patients with depression. *BMC Proc*. 2015;9(S1):A31. doi:10.1186/1753-6561-9-S1-A31.

55. Pence BW, Miller WC, Whetten K, Eron JJ, Gaynes BN. Prevalence of DSM-IV-defined mood, anxiety, and substance use disorders in an HIV clinic in the Southeastern United States. *J Acquir Immune Defic Syndr*. 2006;42(3):298-306. doi:10.1097/01.qai.0000219773.82055.aa.

56. Siu AL. Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: U.S. preventive services task force recommendation statement. *Ann Intern Med*. 2015. doi:10.7326/M15-2023.

57. Wilk AI, Jensen NM, Havighurst TC. Meta-analysis of randomized control trials addressing brief interventions in heavy alcohol drinkers. *J Gen Intern Med*. 1997;12(5):274-283. doi:10.1046/j.1525-1497.1997.012005274.x.

58. Magill M, Ray LA. Cognitive-Behavioral Treatment With Adult Alcohol and Illicit Drug Users: A Meta-Analysis of Randomized Controlled Trials. *J Stud Alcohol Drugs*. 2009;70(4):516-527. doi:10.15288/jsad.2009.70.516.

59. Streeton C, Whelan G. Naltrexone, a relapse prevention maintenance treatment of alcohol dependence: a meta-analysis of randomized controlled trials. *Alcohol Alcohol*. 2001;36(6):544-552. doi:10.1093/alcalc/36.6.544.

60. Ledgerwood DM, Yskes R. Smoking Cessation for People Living With HIV/AIDS: A Literature Review and Synthesis. *Nicotine Tob Res*. 2016;18(12):2177-2184. doi:10.1093/ntr/ntw126.

61. Brown JL, Demartini KS, Sales JM, Swartzendruber AL, Diclemente RJ. Interventions to reduce alcohol use among HIV-Infected individuals: A review and critique of the literature. *Curr HIV/AIDS Rep*. 2013;10(4):356-370. doi:10.1007/s11904-013-0174-8.

62. Westreich D, Edwards JK, Lesko CR, Cole SR, Stuart EA. Target Validity and the Hierarchy of Study Designs. *Am J Epidemiol*. 2019. doi:10.1093/aje/kwy228.

63. Kitahata MM, Rodriguez B, Haubrich R, et al. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. *Int J Epidemiol*. 2008;37(5):948-955. doi:10.1093/ije/dym231.

64. Lesko CR, Todd J V., Cole SR, et al. Mortality under plausible interventions on antiretroviral treatment and depression in HIV-infected women: an application of the parametric g-formula. *Ann Epidemiol*. September 2017:S1047-2797. doi:10.1016/j.annepidem.2017.08.021.

65. Breskin A, Westreich D, Cole SR, et al. The Effects of Hepatitis C Infection and Treatment on All-cause Mortality Among People Living With Human Immunodeficiency Virus. *Clin Infect Dis*. 2019. doi:10.1093/cid/ciy588.

66. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560. doi:10.1097/00001648-200009000-00011.

67. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect. *Math Model*. 1986;7(9-12):1393-1512. doi:10.1016/0270-0255(86)90088-6.

68. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2012;31(18):2000-2009. doi:10.1002/sim.5316.

69. Taubman SL, Robins JM, Mittleman M a., Hernán M a. Intervening on risk factors for coronary heart disease: An application of the parametric g-formula. *Int J Epidemiol*. 2009;38(6):1599-1611. doi:10.1093/ije/dyp192.

70. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: Demonstration of a causal inference technique. *Am J Epidemiol*. 2011;173(7):731-738. doi:10.1093/aje/kwq472.

71. Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology*. 2014;25(6):889-897. doi:10.1097/EDE.0000000000000160.

72. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the G formula. *Am J Epidemiol*. 2013;177(9):989-996. doi:10.1093/aje/kws343.

73. Edwards JK, Cole SR, Westreich D, et al. Age at Entry into Care, Timing of Antiretroviral Therapy Initiation, and 10-Year Mortality among HIV-Seropositive Adults in the United States. *Clin Infect Dis*. 2015;61(7):1189-1195. doi:10.1093/cid/civ463.

74. Lesko CR, Todd J V., Cole SR, et al. Mortality under plausible interventions on antiretroviral treatment and depression in HIV-infected women: An application of the parametric g-formula. *Annals of Epidemiology*. 2017.

75. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative Effectiveness of Dynamic Treatment Regimes: An Application of the Parametric G-Formula. *Stat Biosci*. 2011;3(1):119-143. doi:10.1007/s12561-011-9040-7.

76. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656-664. doi:10.1093/aje/kwn164.

77. Eekhout I, de Boer RM, Twisk JWR, de Vet HCW, Heymans MW. Missing data: a systematic review of how they are reported and handled. *Epidemiology*. 2012;23(5):729-732. doi:10.1097/EDE.0b013e3182576cdb.

78. Young JG, Cain LE, Robins JM, Reilly EJ, Hernán MA. Comparative Effectiveness of Dynamic Treatment Regimes: An Application of the Parametric G-Formula. *Stat Biosci*. 2011;3(1):119-143. doi:10.1007/s12561-011-9040-7.

79. Demissie S, LaValley MP, Horton NJ, Glynn RJ, Cupples LA. Bias due to missing exposure data using complete-case analysis in the proportional hazards regression model. *Stat Med*. 2003;22(4):545-557. doi:10.1002/sim.1340.

80. Breslow NE, Lumley T, Ballantyne CM, Chambless LE, Kulich M. Using the whole cohort in the analysis of case-cohort data. *Am J Epidemiol*. 2009;169(11):1398-1405. doi:10.1093/aje/kwp055.

81. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol*. 1995;142(12):1255-1264. doi:10.1002/ajmg.a.33795.

82. Neyman J. Contribution to the Theory of Sampling Human Populations. *J Am Stat Assoc*. 1938;33(201):101-116. doi:10.1080/01621459.1938.10503378.

83. Hanley JA, Csizmadi I, Collet JP. Two-stage case-control studies: Precision of parameter estimates and considerations in selecting sample size. *Am J Epidemiol*. 2005;162(12):1225-1234. doi:10.1093/aje/kwi340.

84. Breslow NE, Lumley T, Ballantyne CM, Chambless LE, Kulich M. Improved Horvitz–Thompson Estimation of Model Parameters from Two-phase Stratified Samples: Applications in Epidemiology. *Stat Biosci*. 2009;1(1):32-49. doi:10.1007/s12561-009-9001-6.

85. White JE. A two stage design for the study of the relationship between a rare exposure and a rare disease. *Am J Epidemiol*. 1982;115(1):119-128. http://www.ncbi.nlm.nih.gov/pubmed/7055123.

86. Flanders WD, Greenland S. Analytic methods for two-stage case-control studies and other stratified designs. *Stat Med*. 1991;10(5):739-747. doi:10.1002/sim.4780100509.

87. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. *J Am Stat Assoc*. 1994;89(427):846-866. doi:10.1080/01621459.1994.10476818.

88. Hubbard AE, Van Der Laan MJ. Population intervention models in causal inference. *Biometrika*. 2008. doi:10.1093/biomet/asm097.

89. Barkan SE, Melnick SL, Preston-Martin S, et al. The Women's Interagency HIV Study. WIHS Collaborative Study Group. *Epidemiology*. 1998;9(2):117-125. doi:10.1097/00001648-199803000-00004.

90. Bacon MC, Von Wyl V, Alden C, et al. The Women's Interagency HIV Study: an Observational Cohort Brings Clinical Sciences to the Bench. *Clin Diagn Lab Immunol*. 2005;12(9):1013-1019. doi:10.1128/CDLI.12.9.1013–1019.2005.

91. Adimora AA, Ramirez C, Benning L, et al. Cohort Profile: The Women's Interagency HIV Study (WIHS). *Int J Epidemiol*. 2018;47(2):393-394i. http://dx.doi.org/10.1093/ije/dyy021.

92. Feldman JG, Minkoff H, Schneider MF, et al. Association of cigarette smoking with HIV prognosis among women in the HAART era: A report from the women's interagency HIV study. *Am J Public Health*. 2006. doi:10.2105/AJPH.2005.062745.

93. Cook RL, Zhu F, Belnap BH, et al. Longitudinal trends in hazardous alcohol consumption among women with human immunodeficiency virus infection, 1995-2006. *Am J Epidemiol*. 2009. doi:10.1093/aje/kwp004.

94. Theall KP, Clark RA, Powell A, Smith H, Kissinger P. Alcohol consumption, art usage and high-risk sex among women infected with HIV. *AIDS Behav*. 2007. doi:10.1007/s10461-006-9159-6.

95. Reynolds NR. Cigarette smoking and HIV: More evidence for action. In: *AIDS Education and Prevention*. Vol 21. ; 2009:106-121. doi:10.1521/aeap.2009.21.3_supp.106.

96. Breslow NE, Cain KC. Logistic regression for two-stage case-control data. *Biometrika*. 1988;75(1):11-20. doi:10.1093/biomet/75.1.11.

97. Danaei G, Hernan MA, Hu FB. Using parametric G-formula to estimate the effect of multiple lifestyle and dietary interventions for preventing type 2 diabetes in a prospective cohort. *Am J Epidemiol*. 2011;173:S30. doi:http://dx.doi.org/10.1093/aje/kwr181.

98. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the G formula. *Am J Epidemiol*. 2013;177(9):989-996. doi:10.1093/aje/kws343.

99. Edwards JK, Cole SR, Westreich D, et al. Age at Entry Into Care, Timing of Antiretroviral Therapy Initiation, and 10-Year Mortality Among HIV-Seropositive Adults in the United States. *Clin Infect Dis*. 2015;61(7):1189-1195. doi:10.1093/cid/civ463.

100. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology*. 2003;14(6):680-686. doi:10.1097/01.EDE.0000081989.82616.7d.

101. Robins JM, Hernán M a, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560. doi:10.1097/00001648-200009000-00011.

102. Edwards JK, Cole SR, Westreich D, et al. Loss to clinic and five-year mortality among HIV-infected antiretroviral therapy initiators. *PLoS One*. 2014;9(7). doi:10.1371/journal.pone.0102305.

103. Cates JE, Westreich D, Unger HW, et al. Intermittent preventive therapy in pregnancy and incidence of low birth weight in malaria-endemic countries. *Am J Public Health*. 2018. doi:10.2105/AJPH.2017.304251.

104. Young JG, Hernán MA, Robins JM. Identification, estimation and approximation of risk und interventions that depend on the natural value of treatment using observational data. *Epidemiol Method*. 2014. doi:10.1515/em-2012-0001.

105. Robins JM, Hernan MA, Siebert U. Comparative Quantification of Health Risks. *Eff Mult Interv*. 2004:1651-1800.

106. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2012;31(18):2000-2009. doi:10.1002/sim.5316.

107. Edwards JK, Cole SR, Westreich D. All your data are always missing: Incorporating bias due to measurement error into the potential outcomes framework. *Int J Epidemiol*. 2015;44(4):1452-1459. doi:10.1093/ije/dyu272.

108. VanderWeele TJ. Concerning the Consistency Assumption in Causal Inference. *Epidemiology*. 2009;20(6):880-883. doi:10.1097/EDE.0b013e3181bd5638.

109. Pearl J. On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology*. 2010;21(6):872-875. doi:10.1097/EDE.0b013e318209dc0f.

110. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health*. 2006;60(7):578-586. doi:10.1136/jech.2004.029496.

111. Westreich D, Cole SR. Invited commentary: Positivity in practice. *Am J Epidemiol*. 2010;171(6):674-677. doi:10.1093/aje/kwp436.

112. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. *N Engl J Med*. 1997. doi:10.1056/NEJM199709113371101.

113. Ray WA. Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. *Am J Epidemiol*. 2003;158(9):915-920. doi:10.1093/aje/kwg231.

114. AIDSinfo. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV. *Dep Heal Hum Serv*. 2018.

115. Cole SR, Hudgens MG, Brookhart MA, Westreich D. Risk. *Am J Epidemiol*. 2015;181(4):246-250. doi:10.1093/aje/kwv001.

116. Cole SR, Lau B, Eron JJ, et al. Estimation of the standardized risk difference and ratio in a competing risks framework: Application to injection drug use and progression to AIDS after initiation of antiretroviral therapy. *Am J Epidemiol*. 2015. doi:10.1093/aje/kwu122.

117. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: A systematic review and meta-analysis. *BMJ*. 2011. doi:10.1136/bmj.d671.

118. Monto A, Patel K, Bostrom A, et al. Risks of a Range of Alcohol Intake on Hepatitis C-Related Fibrosis. *Hepatology*. 2004. doi:10.1002/hep.20127.

119. U.S. Department of Health and Human Services. Helping patients who drink too much: A clinicians guide. National Institute on Alcohol Abuse and Alcoholism. http://www.niaaa.nih.gov/publications. Published 2005.

120. Dieperink E, Ho SB, Heit S, Durfee JM, Thuras P, Willenbring ML. Significant reductions in drinking following brief alcohol treatment provided in a hepatitis C clinic. *Psychosomatics*. 2010. doi:10.1176/appi.psy.51.2.149.

121. Hoffman AM, Redding CA, Goldberg D, et al. Computer expert systems for African-American smokers in physicians offices: A feasibility study. *Prev Med (Baltim)*. 2006. doi:10.1016/j.ypmed.2006.03.025.

122. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999. doi:10.1097/00001648-199901000-00008.

123. Brookhart MA. Counterpoint: The Treatment Decision Design. *Am J Epidemiol*. 2015;182(10):840-845. doi:10.1093/aje/kwv214.

124. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Appl Psychol Meas*. 1977. doi:10.1177/014662167700100306.

125. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc*. 1958;53(282):457-481. doi:10.2307/2281868.

126. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. *Chapman Hall, New York*. 1993. doi:10.1111/1467-9639.00050.

127. Cunningham CO. Opioids and HIV infection: From pain management to addiction treatment. *Top Antivir Med*. 2017.

128. Deyo RA, Hallvik SE, Hildebran C, et al. Association Between Initial Opioid Prescribing Patterns and Subsequent Long-Term Use Among Opioid-Naïve Patients: A Statewide Retrospective Cohort Study. *J Gen Intern Med*. 2017. doi:10.1007/s11606-016-3810-3.

129. Schieber LZ, Guy GP, Seth P, et al. Trends and Patterns of Geographic Variation in Opioid Prescribing Practices by State, United States, 2006-2017. *JAMA Netw open*. 2019. doi:10.1001/jamanetworkopen.2019.0665.

130. Josephs JS, Fleishman JA, Korthuis PT, Moore RD, Gebo KA. Emergency department utilization among HIV-infected patients in a multisite multistate study. *HIV Med*. 2010;11(1):74-84. doi:10.1111/j.1468-1293.2009.00748.x.

131. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res*. 2011. doi:10.1177/0962280210395740.

132. Firth D. Bias Reduction of Maximum Likelihood. *Biometrika*. 1993. doi:10.1093/biomet/80.1.27.

133. Galvan FH, Bing EG, Fleishman JA, et al. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: Results from the HIV cost and services utilization study. *J Stud Alcohol*. 2002. doi:10.15288/jsa.2002.63.179.

134. Duprez DA, Neuhaus J, Kuller LH, et al. Inflammation, coagulation and cardiovascular disease in HIV-infected individuals. *PLoS One*. 2012;7(9). doi:10.1371/journal.pone.0044454.

135. Hernán M, Robins J. Estimating causal effects from epidemiological data. *Epidemiol Community Heal*. 2006;60(7):578-586. doi:10.1136/jech.2004.029496.

136. Cropsey KL, Willig JH, Mugavero MJ, et al. Cigarette Smokers are Less Likely to Have Undetectable Viral Loads: Results from Four HIV Clinics. *J Addict Med*. 2016. doi:10.1097/ADM.0000000000000172.

137. Chander G, Lau B, Moore RD. Hazardous alcohol use: A risk factor for non-adherence and lack of suppression in HIV infection. *J Acquir Immune Defic Syndr*. 2006. doi:10.1097/01.qai.0000243121.44659.a4.

138. Samet JH, Horton NJ, Meli S, Freedberg KA, Palepu A. Alcohol Consumption and Antiretroviral Adherence among HIV-Infected Persons with Alcohol Problems. *Alcohol Clin Exp Res*. 2004. doi:10.1097/01.ALC.0000122103.74491.78.

139. Kozak MS, Mugavero MJ, Ye J, et al. Patient reported outcomes in routine care: Advancing data capture for HIV cohort research. *Clin Infect Dis*. 2012;54(1):141-147. doi:10.1093/cid/cir727.

140. Moore D, Aveyard P, Connock M, Wang D, Fry-Smith A, Barton P. Effectiveness and safety of nicotine replacement therapy assisted reduction to stop smoking: Systematic review and meta-analysis. *BMJ*. 2009. doi:10.1136/bmj.b1024.

141. Brown JL, Demartini KS, Sales JM, Swartzendruber AL, Diclemente RJ. Interventions to reduce alcohol use among HIV-Infected individuals: A review and critique of the literature. *Curr HIV/AIDS Rep*. 2013. doi:10.1007/s11904-013-0174-8.

142. Samet JH, Walley AY. Interventions targeting HIV-infected risky drinkers: Drops in the bottle. *Alcohol Res Heal*. 2010.

143. Samet JH, Horton NJ, Meli S, et al. A randomized controlled trial to enhance antiretroviral therapy adherence in patients with a history of alcohol problems. *Antivir Ther*. 2005.

144. Pearl J. On the consistency rule in causal inference: Axiom, definition, assumption, or theorem? *Epidemiology*. 2010;21(6):872-875. doi:10.1097/EDE.0b013e3181f5d3fd.

145. Robins JM, Hernán M a., Rotnitzky A. Invited commentary: Effect modification by time-varying covariates. *Am J Epidemiol*. 2007;166(9):994-1002. doi:10.1093/aje/kwm231.

146. Lash TL, Schisterman EF. New Designs for New Epidemiology. *Epidemiology*. 2017;29(1):2017-2018. doi:10.1097/EDE.0000000000000768.

147. Daniel RM, Kenward MG, Cousens SN, De Stavola BL. Using causal diagrams to guide analysis in missing data problems. In: *Statistical Methods in Medical Research*. ; 2012. doi:10.1177/0962280210394469.

148. Hernán MA. Invited commentary: Selection bias without colliders. *Am J Epidemiol*. 2017. doi:10.1093/aje/kwx077.

149. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing Study Results: A Potential Outcomes Perspective. *Epidemiology*. 2017. doi:10.1097/EDE.0000000000000664.

150. Stuart EA, Bradshaw CP, Leaf PJ. Assessing the generalizability of randomized trial results to target populations. *Prev Sci*. 2015. doi:10.1007/s11121-014-0513-z.

151. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *Am J Epidemiol*. 2010. doi:10.1093/aje/kwq084.

152. Rudolph KE, Díaz I, Rosenblum M, Stuart EA. Estimating population treatment effects from a survey subsample. *Am J Epidemiol*. 2014. doi:10.1093/aje/kwu197.

153. Hoenigl M, Chaillon A, Moore DJ, et al. Rapid HIV Viral Load Suppression in those Initiating Antiretroviral Therapy at First Visit after HIV Diagnosis. *Sci Rep*. 2016. doi:10.1038/srep32947.

154. Pilcher CD, Ospina-Norvell C, Dasgupta A, et al. The Effect of Same-Day Observed Initiation of Antiretroviral Therapy on HIV Viral Load and Treatment Outcomes in a US Public Health Setting. *J Acquir Immune Defic Syndr*. 2017. doi:10.1097/QAI.0000000000001134.

155. Ford N, Migone C, Calmy A, et al. Benefits and risks of rapid initiation of antiretroviral therapy. *AIDS*. 2018. doi:10.1097/QAD.0000000000001671.

156. Sun B, Perkins NJ, Cole SR, et al. Inverse-Probability-Weighted Estimation for Monotone and Nonmonotone Missing Data. *Am J Epidemiol*. 2018. doi:10.1093/aje/kwx350.

157. Sun BL, Tchetgen Tchetgen EJ. On Inverse Probability Weighting for Nonmonotone Missing at Random Data. *J Am Stat Assoc*. 2018. doi:10.1080/01621459.2016.1256814.

158. The Antiretroviral Therapy Cohort Collaboration. Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996-2006: collaborative analysis of 13 HIV cohort studies. *Clin Infect Dis*. 2010;50(10):1387-1396. doi:10.1086/652283.

159. French AL, Gawel SH, Hershow R, et al. Trends in mortality and causes of death among women with HIV in the United States: a 10-year study. *J Acquir Immune Defic Syndr*. 2009;51(4):399-406. doi:10.1097/QAI.0b013e3181acb4e5.

160. Weber R, Ruppik M, Rickenbach M, et al. Decreasing mortality and changing patterns of causes of death in the Swiss HIV Cohort Study. *HIV Med*. 2013;14(4):195-207. doi:10.1111/j.1468-1293.2012.01051.x.

161. Justice AC, Dombrowski E, Conigliaro J, et al. Veterans Aging Cohort Study (VACS): Overview and description. *Med Care*. 2006;44(8 Suppl 2):S13-24. doi:10.1097/01.mlr.0000223741.02074.66.

162. Danaei G, Pan A, Hu FB, Hernán MA. Hypothetical Midlife Interventions in Women and Risk of Type 2 Diabetes. *Epidemiology*. 2013;24(1):122-128. doi:10.1097/EDE.0b013e318276c98a.

163. Egger M, Hirschel B, Francioli P, et al. Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. Swiss HIV Cohort Study. *BMJ*. 1997;315(7117):1194-1199. doi:10.1136/bmj.315.7117.1194.

164. Baker J V., Sharma S, Achhra AC, et al. Changes in Cardiovascular Disease Risk Factors With Immediate Versus Deferred Antiretroviral Therapy Initiation Among HIV-Positive Participants in the START (Strategic Timing of Antiretroviral Treatment) Trial. *J Am Heart Assoc*. 2017. doi:10.1161/JAHA.116.004987.

165. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol*. 1995;142(12):1255-1264. doi:10.1093/aje/kwi188.