## Title

Novel Methods for Bias Analysis of Time-Varying Uncontrolled Confounding in Epidemiologic Studies of Chronic Diseases



## Author

Soohoo, Melissa Anne

## Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Novel Methods for Bias Analysis of Time-Varying Uncontrolled Confounding in Epidemiologic Studies of Chronic Diseases

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Epidemiology

by

Melissa Anne Soohoo

2022

ABSTRACT OF THE DISSERTATION

Novel Methods for Bias Analysis of Time-Varying Uncontrolled Confounding in Epidemiologic Studies of Chronic Diseases

by

Melissa Anne Soohoo

Doctor of Philosophy in Epidemiology

University of California, Los Angeles, 2022

Professor Onyebuchi Aniweta Arah, Chair

Observational studies investigating chronic diseases are valuable in informing public health guidelines. Yet, it is pertinent that studies are evaluated for threats of biases including uncontrolled confounding, which may produce misguided conclusions if left unaddressed. The fitting of marginal structural models using g-computation is a method capable in evaluating time-varying study designs but methods in addressing the bias due to time-varying uncontrolled confounding are understudied. This dissertation graphically describes the types of time-varying uncontrolled confounding structures. Combined with g-computation, the dissertation highlights the development of two novel bias analysis methods to address time-varying uncontrolled confounders in a plasmode simulation framework. The first method utilizes a bias offset to subtract confounding from the outcome and formally is applied to an investigation of elevated depressive symptoms with cancer outcomes. The last method employs an inverse probability of

uncontrolled confounder weight as the bias weighting method and is applied to a study of elevated depressive symptoms with cardiovascular disease. Several settings can describe time-varying uncontrolled confounding including relations into subsequently measured confounders, exposures, and the outcome. Simulated results demonstrated the most biased estimates when time-varying uncontrolled confounding were consistently strong over time. In simulations, both the bias offset and bias weighting methods could recover a true estimate from mis-specified, biased models across a series of different bias parameters and effect estimates. In applied illustrations, null and modest relationships were observed between elevated depressive symptoms and incidence of cancer outcomes and cardiovascular disease, respectively. For all illustrations, applied bias analysis suggested robust results to modest levels of confounding. Time-varying uncontrolled confounding can immensely impact observed estimates. This dissertation provides a principled approach to alternative explanations (due to mixing of effects) to enable more credible causal inference in the health sciences. Applied examples demonstrate two novel bias analysis methods and a guided framework of how bias analysis can be combined with causal inference methods. Longitudinal observational studies are integral in advising public health and thus the impact of sources of bias warrants more recognition and careful evaluation using bias analysis methods.

The dissertation of Melissa Anne Soohoo is approved.

Elizabeth Rose Mayeda

Roch Arnaud Kibsa Nianogo

May-Choo Wang

Zuofeng Zhang

Onyebuchi Aniweta Arah, Committee Chair

University of California, Los Angeles

2022

## Table of Contents

**List of Figures** ..... viii

**List of Tables** ..... x

**List of Abbreviations** ..... xii

**Acknowledgements** ..... xiii

**1. Introduction** ..... 1

1.1 Chronic Disease Epidemiology ..... 1

1.1.1 Depressive Symptoms and Cancer ..... 2

1.1.2 Depressive Symptoms and Cardiovascular Disease ..... 3

1.2 Causal Inference and Assumptions in a Time-Varying Setting ..... 5

1.3 Bias Analysis ..... 8

1.4 Dissertation Structure and Specific Aims ..... 11

**2. Study 1: Time-Varying Uncontrolled Confounding in Cohort Data Analyzed Using G-Computation: Simulated Applications to Cancer Epidemiology** ..... 14

2.1 Abstract ..... 15

2.2 Background ..... 17

2.2.1 Time-Varying Uncontrolled Confounding Structures ..... 18

2.3 Methods ..... 22

2.3.1 Hypothetical Illustrations and Motivation in Cancer Epidemiology ..... 27

2.4 Results ..... 28

2.5 Discussion ..... 33

2.6 Appendix ..... 37

**3. Study 2: Investigating Whether Elevated Depressive Symptoms Affect Cancer Outcomes: G-Computation with Bias Analysis to Offset Time-Varying Uncontrolled Confounding** ..... 41

3.1 Abstract ..... 42

3.2 Introduction ..... 44

3.3 Methods ..... 45

3.3.1 Illustration with the Health and Retirement Study ..... 52

3.4 Results ..... 56

3.4.1 Simulations Demonstrating the Bias Offset ..... 56

3.4.2 Clinical Illustrations with the Health and Retirement Study ..... 61

3.5 Discussion ..... 65

3.6 Appendix ..... 69

**4. Study 3: Inverse Probability Weighted G-Computation Bias Analysis for Time-Varying Uncontrolled Confounding: Applied to an Investigation of the Link Between Elevated Depressive Symptoms and Cardiovascular Disease** ..... 75

4.1 Abstract ..... 76

4.2 Introduction ..... 78

4.3 Methods ..... 79

4.3.1 Illustration with the Health and Retirement Study ..... 87

4.4 Results ..... 91

4.4.1 Simulations Demonstrating the Bias Weight ..... 91

4.4.2 Clinical Illustration with the Health and Retirement Study ..... 97

4.5 Discussion ..... 105

4.6 Appendix ..... 111

**5. Conclusion** ..... 118

**References** ..... 122

## List of Figures

**Figure 1.1.** Relationship of the Specific Aims for this Dissertation ..... 13

**Figure 2.1.** Directed Acyclic Graphs Depicting the Data-Generating Mechanism and Types of Uncontrolled Confounding of the Joint Effect of $X_t$ and Y over Four Time Points..... 20

**Figure 2.2.** Simulated Illustration of the Joint Odds Ratio of Overweight and Obesity Status with Related Cancers over Follow-up Across Multiple Biasing Scenarios in Simulation 1.2 ..... 32

**Appendix Figure 2.1.** Directed Acyclic Graph Depicting the Data-Generating Mechanism and Uncontrolled Confounding of the Joint Association of $X_t$ and Y over Two Time Points..... 37

**Appendix Figure 2.2.** Simulated Illustration of the Joint Mean Difference in LDL from Statin Non-Adherence Before and During Breast Cancer Treatment in Simulation 1.1 ..... 38

**Figure 3.1.** Directed Acyclic Graphs of Time-Varying Exposures Depicting A) Only One Uncontrolled Confounder, B) Only Two Time-Varying Uncontrolled Confounders and C) Two Time-Varying Uncontrolled Confounders and an Unmeasured Mediator ..... 47

**Figure 3.2.** Directed Acyclic Graphs Depicting A) Pre-Intervention, B) Partial-Intervention after use of the Bias Offset and C) Post-Intervention after G-Computation..... 50

**Figure 3.3.** Directed Acyclic Graphs for the Health and Retirement Study Assessing the Joint Effect of Elevated Depressive Symptoms and Cancer Incidence A) with and B) without Time-Varying Uncontrolled Confounders..... 54

**Appendix Figure 3.1.** Cohort Construction for Health and Retirement Study Analyses..... 69

**Figure 4.1.** Directed Acyclic Graphs of Time-Varying Exposures Depicting A) Only Two Time-Varying Uncontrolled Confounders and B) Two Time-Varying Uncontrolled Confounders and an Unmeasured Mediator..... 82

**Figure 4.2.** Directed Acyclic Graphs Depicting A) Pre-Intervention, B) Partial-Intervention After Inverse Probability of Uncontrolled Confounder Weight, and C) the Post-Intervention after G-computation ..... 85

**Figure 4.3.** Bias Analysis Where $U_5$ is a Mediator and Confounder for the A) Joint Relative Risk and B) Joint Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants ..... 103

**Figure 4.4.** Bias Analysis Where $U_5$ is a Confounder only for the A) Joint Relative Risk and B) Joint Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants..... 104

**Appendix Figure 4.1.** Cohort Construction for Health and Retirement Study Analyses..... 111

**Appendix Figure 4.2.** Directed Acyclic Graphs for the Health and Retirement Study Assessing the Joint Effect of Elevated Depressive Symptoms and Incident CVD A) with and B) without Time-Varying Uncontrolled Confounders..... 112

**Appendix Figure 4.3.** Bias Analysis for the Joint A) Relative Risk and B) Risk Difference of Elevated Depressive Symptoms and Incident CVD over Maximum Follow-up in the Health and Retirement Study Participants Where $U_5$ is a i) Mediator and Confounder or ii) a Confounder only ..... 113

**Appendix Figure 4.4.** Bias Analysis for the Joint A) Relative Risk and B) Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants, when Changing the $U_5$ Parameters**.** ..... 115

## List of Tables

**Table 1.1** Summary of Assumptions Required for Causal Inference ..... 7

**Table 2.1.** Parameters for A) Simulation 1.1 for Analyses of Two Time Points and B) Simulation 1.2 for Analyses of Four Time Points ..... 23

**Table 2.2.** Regression Parameters and the Joint Mean Difference in LDL from Statin Non-Adherence Before and During Breast Cancer Treatment in Simulation 1.1 ..... 30

**Appendix Table 2.1.** Joint Odds Ratio Over Follow-up Across Multiple Confounding Scenarios ..... 39

**Appendix Table 2.2.** Joint Odds Ratio of Overweight and Obesity Status with Related Cancers over Follow-up Across Multiple Biasing Scenarios ..... 40

**Table 3.1.** Parameters for Simulations 2.1-2.3 ..... 46

**Table 3.2.** Simulation Results for the Joint Risk Difference in Simulation 2.1 ..... 57

**Table 3.3.** Simulation Results for the Joint Risk Difference in Simulation 2.2 ..... 59

**Table 3.4.** Simulation Results for the Joint Risk Differences in Simulation 2.3 ..... 60

**Table 3.5.** Cohort Descriptives of 12,502 Health and Retirement Study Participants Stratified by Baseline Level of Depressive Symptoms ..... 62

**Table 3.6.** Illustrative Example Investigating the Relationship of Elevated Depressive Symptoms (CES-D≥4) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior ..... 64

**Appendix Table 3.1.** Cohort Descriptives of 12,502 Health and Retirement Study Participants Stratified by Elevated Depressive Symptoms Across Waves 4 and 5 ..... 70

**Appendix Table 3.2.** Sensitivity Analyses for Relationship of Elevated Depressive Symptoms (CES-D≥4) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying

Uncontrolled Confounding due to Sedentary Behavior Across Multiple G-computation Model Specifications ..... 72

**Appendix Table 3.3.** Sensitivity Analyses Investigating the Relationship of Elevated Depressive Symptoms (CES-D≥3) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior ..... 73

**Appendix Table 3.4.** Sensitivity Analyses for Relationship of Elevated Depressive Symptoms (CES-D≥3) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior Across Multiple G-computation Model Specifications ..... 74

**Table 4.1.** Parameters for Simulations 3.1-3.3 ..... 81

**Table 4.2.** Simulation Results for the Joint Risk Difference in Simulation 3.1 that Included $Z_t$ and Excluded $Z_t$ in the IPUW ..... 92

**Table 4.3.** Simulation Results for the Joint Risk Difference in Simulation 3.2 that Included $Z_t$ and Excluded $Z_t$ in the IPUW ..... 93

**Table 4.4.** Simulation Results for the Joint Odds Ratio in Simulation 3.3 that Included $Z_t$ and Excluded $Z_t$ in the IPUW ..... 95

**Table 4.5.** Cohort Descriptives of 10,212 Health and Retirement Study Participants Stratified by Baseline Level of Depressive Symptoms ..... 98

**Table 4.6.** Joint Relationship of Elevated Depressive Symptoms (CES-D ≥4) and Incident CVD over Ten Years in 10,212 Health and Retirement Study Participants and Under Sensitivity Analyses ..... 101

**Appendix Table 4.1.** Cohort Descriptives of 10,212 Health and Retirement Study Participants Stratified by Elevated Depressive Symptoms Across Waves 4 and 5 ..... 116

## List of Abbreviations

| | |
|---|---|
| CES-D | Center for Epidemiological Studies Depression Scale |
| CHD | Coronary Heart Disease |
| CI | Confidence Interval |
| CVD | Cardiovascular Disease |
| DAG | Directed Acyclic Graph |
| DRE | Doubly robust estimation |
| expit | Inverse logit |
| HR | Hazard Ratio |
| HRS | Health and Retirement Study |
| IPTW | Inverse probability treatment weight |
| IPUW | Inverse probability of uncontrolled confounder weight |
| IQR | Interquartile Range |
| LDL | Low-density lipoprotein cholesterol |
| MD | Mean Difference |
| MSM | Marginal structural models |
| OR | Odds Ratio |
| RD | Risk Difference |
| RMSE | Root mean square error |
| RR | Relative Risk |
| SD | Standard Deviation |
| SI | Simulation Interval |

**Acknowledgements**

First and foremost, I am immensely grateful for Dr. Arah. His superb mentorship, expertise, and unwavering support have been invaluable. It has been a great privilege to work with you on my most passionate projects. Thank you for being my biggest champion. I am also appreciative of the support provided by my dissertation committee, along with their boundless guidance, enthusiasm, and encouragement. Thank you to my HSC colleagues who have inspired and fostered much of my work and honed my skills. Thank you to the fellowships and teaching apprenticeships from the NIH/NCI Cancer Epidemiology Training Program (T32CA009142), UCLA Graduate Division, and UCLA Department of Epidemiology for the funding and support for this dissertation and doctorate.

And a very special thanks to my colleagues and classmates who have become vital confidants including Adrian Bacong, Kosuke Inoue, Allie Rosen, and Yongfu Yu. Thank you, James Li for your unconditional support and having my back, literally for each day of this dissertation. Finally, none of this would have been possible without the infinite support of my parents, siblings, and extended family for which I am eternally grateful. Thank you for everything.

# VITA

## EDUCATION

| | |
|---|---|
| 2014 | MPH in Epidemiology<br>University of California, Irvine, CA, USA |
| 2012 | BS in Molecular, Cell and Developmental Biology<br>University of California, Los Angeles, CA, USA |

## RESEARCH EXPERIENCE

| | |
|---|---|
| 2014-2021 | Data Scientist/Epidemiology Methodologist/Assistant Statistician<br>Southern California Institute for Research and Education,<br>Tibor Rubin VA Medical Center<br>Long Beach, CA, USA &<br><br>Division of Nephrology, Hypertension & Kidney Transplantation<br>University of California, Irvine<br>Irvine, CA, USA |
| 2020 | Summer Associate<br>Analysis Group<br>Los Angeles, CA, USA |

## TEACHING EXPERIENCE

| | |
|---|---|
| Winter 2021 | Statistical Methods for Epidemiology |
| Fall 2019, Winter 2020, Fall 2020 | Computer Management and Analysis of Health Data Using SAS |
| Spring 2020 | Advanced SAS Techniques for Management and Analysis of Epidemiologic Data |
| Spring 2020 | Applied Epidemiologic Analysis |
| Winter 2014 | Introduction to Epidemiology |

## SELECT PEER-REVIEWED PUBLICATIONS

**Soohoo M**, Obi Y, Rivara MB, Adams SV, Lau WL, Rhee CM, Kovesdy CP, Kalantar-Zadeh K, Arah OA, Mehrotra R, Streja E. Comparative Effectiveness of Dialysis Modality on Laboratory Parameters of Mineral Metabolism. Am J Nephrol. 2022:1-12. Epub 20220228. doi: 10.1159/000521508. PubMed PMID: 35226895.

**Soohoo M,** Streja E, Hsiung JT, Kovesdy CP, Kalantar-Zadeh K, Arah OA. Cohort Study and Bias Analysis of the Obesity Paradox Across Stages of Chronic Kidney Disease. J Ren Nutr. 2021. Epub 20211130. doi: 10.1053/j.jrn.2021.10.007. PubMed PMID: 34861399.

**Soohoo M**, Hashemi L, Hsiung JT, Moradi H, Budoff MJ, Kovesdy CP, Kalantar-Zadeh K, Streja E. Risk of Atherosclerotic Cardiovascular Disease and Nonatherosclerotic Cardiovascular Disease Hospitalizations for Triglycerides Across Chronic Kidney Disease Stages Among 2.9 Million US Veterans. J Am Heart Assoc. 2021:e022988. Epub 20211103. doi: 10.1161/JAHA.121.022988. PubMed PMID: 34729994.

Yu Y, **Soohoo M**, Sørensen HT, Li J, Arah OA. Gestational Diabetes Mellitus and the Risks of Overall and Type-Specific Cardiovascular Diseases: A Population- and Sibling-Matched Cohort Study. Diabetes Care. 2021. Epub 20211111. doi: 10.2337/dc21-1018. PubMed PMID: 34764208.

# 1. Introduction

## 1.1 Chronic Disease Epidemiology

Cardiovascular disease (CVD) and cancer are the leading causes of morbidity and mortality in developed countries[1,2]. Moreover, mental health disorders including major depression are in the top causes of years lived with disability. Together, these major comorbidities represent the foremost causes of disability-adjusted life years. As the pillars of chronic disease epidemiology, they also have complex relationships with other comorbidities and all-cause mortality. Their presence is a burden not only to the individual but also to the population because treatments or prevention interventions for chronic diseases result in the increased use of health care resources. The total direct cost for CVD and cancer in the US are 294 and 98.6 billion dollars, respectively[3]. Collectively, these risk factors are responsible for over a trillion dollars' worth of indirect costs stemming from lower productivity and disability. Historically, the US has had the greatest healthcare spending among developed countries, yet one of the lowest life expectancies, suggesting that despite these spending efforts the population has yet to experience the expected health benefits[2,4].

Major task forces rely on observational studies to inform policy and screening recommendations in the prevention of adverse events and risk factor identification[5,6]. Yet the presence of systematic biases and errors may inadvertently lead to misguided conclusions. Thus, it is important that research is rigorously tested for biases to appropriately assess the burden of disease and efficiently allocate healthcare costs to improve our population's health. Here, we explore several relationships within chronic disease epidemiology.

### 1.1.1 Depressive Symptoms and Cancer

Depression is one of the most prevalent comorbidities in the world[7]. In the US, it is estimated that 12% of older adults aged 65 years and older have experienced recent mild depressive symptoms[8]. Cancer is one of the leading causes of morbidity in the world, resulting in 10 million deaths and 23.6 million incident cases (17.2 million excluding non-melanoma skin cancer) in 2019[9]. Depression and its symptoms as risk factors for cancer incidence have been postulated for years, yet the epidemiological evidence is relatively mixed[10]. Meta-analyses suggest a small effect between depression and cancer incidence, while individual cohort studies suggest associations in both directions[10-15]. A prospective study demonstrated a strong positive relationship between major depression and cancer risk[10]. Conversely, another study concluded that their data did not support a depression-cancer relationship (standardized incidence ratio [95% confidence interval, CI]: 1.00 [0.97, 1.03] and 0.99 [0.95, 1.02]) after 1-9 and ≥10 years of follow-up, respectively)[13]. These relationships have also been investigated for site-specific cancer. Breast cancer is the leading site-specific cancer for females in the world resulting in almost 2 million cases and an age-standardized incidence rate [95%CI] of 45.9 [41.9, 49.8] per 100,000 person-years[9]. Large individual studies did not observe a relationship between depression and breast cancer[12,14]. Meta-analyses, however, have found positive relationships between depressive symptoms and breast cancer, with the relationship magnified in studies with at least a decade of follow-up (relative risk, RR [95%CI]: 2.50 [1.06, 5.91])[15,16]. Among males, the relationship between depressive symptoms and prostate cancer is less clear and understudied. Globally in 2019, there have been 1.4 million incident prostate cancer cases and an age-standardized incidence rate [95%CI] of 38.6 [33.6, 49.8] per 100,000 person-years[9]. Studies showed that minor depression, a single episode of major depression, and the combination of

depression and anxiety in a meta-analysis were related to elevated prostate cancer risk[10,12,17]. However, a Mendelian randomization did not observe a relationship of genetic markers and prostate cancer after a series of analytical methods[18]. Meta-analyses similarly showed a null or only a small risk albeit subject to several limitations[11,15].

Possible reasons for differences across studies have been varying lengths of follow-up, measurement error of baseline depression due to the range of diagnosis methods, and potential uncontrolled confounding from lifestyle factors including alcohol use, smoking use, and hormones[11,12,15,19,20]. In addition, the suggested mechanistic pathways between depression and cancer incidence generally involve a reduced immune system, circulating pro-inflammatory markers, inhibition of DNA repair or confounding by latent psychological risk factors often unavailable in studies[10]. While depression is a chronic condition, its symptoms may also be transient, therefore providing the framework to examine time-updated relationships[19]. Information on chronic or recurrent depressive symptoms have been limited. A study suggested that two instances of elevated depressive symptoms had a lower hazard ratio (HR) for overall cancer incidence, yet three instances of elevated depressive symptoms demonstrated had a small to null relationship (HR [95%CI]: 1.05 [0.73, 1.52])[21]. Data on site-specific cancers have been equally mixed, where recurrent symptoms were associated with a HR [95%CI] of 1.32 [0.67, 2.60] and 0.60 [0.38, 0.96], respectively for outcomes of breast and prostate cancer incidence[22]. Thus, the relationship of depression and overall cancer or specific cancer sites appears to be unclear and prior study limitations prompt further research in this area.

### 1.1.2 Depressive Symptoms and Cardiovascular Disease

Globally, the prevalence of CVD has been increasing over time, afflicting an estimated 523 million persons[23]. Specifically, total CVD is also described in the context of specific causes,

where coronary heart disease (CHD) and stroke are leading reasons. Between 2005-2014, close to 720,000 incident CHD events occurred in US adults[24]. The overwhelming majority of incident CHD events are due to myocardial infarctions, where a new event occurs every 40 seconds in the US[24]. These cardiac events have been well-studied and noted anxiety and depression as risk factors[25]. Broadly, the field has examined two prevailing research questions in regards to total CVD and its specific causes: the association of depression with incident CVD events, and the association of depression with CVD progression, or CVD event recurrence among prevalent CVD patients[26]. Notably, several large studies have demonstrated that baseline depression and standard deviation increases in logged scores were all associated with greater risks of total CVD and CHD, even after controlling for several lifestyle behaviors and laboratory measurements[27-29]. Multiple meta-analyses have suggested strong positive relationships with depressive symptoms (including major depressive disorder) and risk of CVD subtypes[30-33]. With the body of observational evidence, the American Heart Association campaigned for depression screening in prevalent CHD patients[5].

There is sparse data on depression as a time-varying exposure with CVD related outcomes, even though depression and its symptoms are subject to change and relapse[34,35]. The majority of evidence have been based on baseline study designs. One study suggested that time-varying depressive symptoms were associated with a higher risk of CVD death (HR [95%CI]: 1.30 [1.04, 1.63]), yet only a small association with incident coronary heart disease (HR [95%CI]: 1.11 [0.89, 1.38]) over a median follow-up of almost seven years[36]. Other studies have focused on chronic depressive symptoms but with short-term associations with causes of CVD. These studies have suggested that persistent elevated depressive symptoms were associated with two to three times the risks of the referent group for stroke outcomes[37,38]. The relationships of

persistent depressive symptoms on total CVD outcomes require more research. These results, prior limitations of lack of data on related lifestyle factors and latent psychological factors, and mixed evidence from early clinical trials on the effectiveness of depression treatment on event reductions among CVD patients highlight the need for more research in the field[26,36,39-42]. Given the importance of maintaining cardiovascular health, it is critical to fully understand these relationships and robustness to potential biases.

### 1.2 Causal Inference and Assumptions in a Time-Varying Setting

In observational studies, systematic threats to internal and external validity include confounding, information bias and selection bias[43]. Confounding is said to exist when an extraneous factor distorts the observed exposure-outcome relationship from the counterfactual, or ideal effect had the exposure not been present. Information bias occurs when an exposure and/or outcome are mismeasured, and thus incorrectly classified. Finally, selection bias occurs when the observed exposure-outcome relationship among those who participated in the study is different than those who were eligible for analyses, including non-participants. Each bias necessitates control by study design or analysis; if left unaddressed, their presence can result in a biased effect estimate[43]. Modern methods have been developed for the quantitative bias analysis of these (sources of) bias, where these techniques have laid the foundation of epidemiologic research, including the use of sensitivity analyses to evaluate their impact on a study[43].

Bias requires control in a single time point, or baseline analysis, and create added complexity when examining in a longitudinal or time-varying setting. Such time-varying analyses comprise serial measurements of the exposure and/or outcomes and confounders over distinct time points over the follow-up period. These time-varying exposures can be evaluated as joint or cumulative exposure effects in relation to the outcome. The causal relationships between

exposures, outcomes and covariates are often depicted in directed acyclic graphs (DAGs). In the DAG, variables (or nodes): exposures ($X_t$), outcomes ($Y_t$), and covariates ($Z_t$), are connected with directed arrows (or edges), representing the causal structure and indexed by time t.
A path could consist of multiple variables in a single direction where those in the middle are considered as mediators. A path is blocked when arrowheads meet and result in a collider. Conditioning on a collider, or its descendent will open the path, resulting in collider-stratification bias. An open backdoor path is characterized as a fork, where a third variable has arrows pointing into two other variables and represents a confounder. DAGs for time-varying exposure studies are complicated given the involvement of a time-varying covariate that acts as both a confounder and mediator on different paths. These dueling roles within the causal structure requires opposite actions for bias reduction and as a result, conventional regression cannot adequately estimate the joint effect in this context.

To address time-varying structures and their complexities, the works of Robins et al. and others have championed solutions via the use of inverse probability treatment weighting (IPTW), and the g-computation formula to represent the fitting of marginal structural models (MSM)[44]. Both methods tackle the consequences arising from time-varying confounding and mediation and have grown in applications as technical guides and large longitudinal datasets have increased in availability[45,46]. G-computation offers advantages over IPTW as it is more statistically efficient and can evaluate time-varying interactions between time-varying exposures and time-varying covariates, yet its limitations include the need for high computing capabilities, and the possibility of a g-null paradox[45]. G-computation is increasingly popular because of its flexibility when used in modeling potential outcomes (i.e., outcomes that would have occurred had the exposure been

set to a particular intervention value, perhaps counterfactually) of dynamic treatments, with time-varying confounding[45].

Causal modeling requires several effect identifiability conditions or assumptions and are briefly summarized in Table 1.1[44,45].

**Table 1.1 Summary of Assumptions Required for Causal Inference**

| Causal Assumption | Description |
|---|---|
| Conditional exchangeability | Conditional on Z covariates sufficient to close backdoor paths, the exposure X is independent of the potential outcome, $Y_x$. This is also referred to as the assumption of 'no uncontrolled confounding'. |
| Consistency | Among those with observed X = x, their potential outcome $Y_x$ is their observed outcome Y. |
| Positivity | Whenever the probability of Z is non-null, the probability of X = x should also be non-null for all Z = z. That is, there is no deterministic exposure assignment for any Z covariates deemed sufficient for control of confounding. |
| Well defined intervention | A clear and sufficient definition of treatment such that a unique potential outcome could uniquely exist. |
| No interference | There are no spillover effects, i.e., setting the exposure of one person does not affect the potential outcome of another. If interference exists, then appropriate spill-over effect modeling methods should be used. |
| No other sources of bias | There are no other biases such as measurement error, selection bias, model misspecification and so on. |

One notable assumption is conditional exchangeability, which assumes that the potential outcomes in the exposed and unexposed groups are comparable within strata of covariates Z. This assumption is also referred to as ignorability, exogeneity, or no uncontrolled confounding. In other words, those who were exposed would have the same potential outcome on average as those who were unexposed, if the exposed persons would have been in fact unexposed. This assumption does not only refer to a baseline analysis and extends to include the absence of uncontrolled confounding in a time-varying setting[44]. Specifically, under a time-varying study, this assumption is interpreted as given the prior history of the exposure and covariates, the potential outcome is independent of the current exposure level that is actually received.

All assumptions, including the presence of conditional exchangeability, can be strong and may be untestable. Confounder selection is based on theoretical and field expertise to select a sufficient set of confounders, yet data availability may result in a single or even multiple unmeasured and uncontrolled time-varying confounders excluded from analyses. Moreover, it may be implausible to assume that wholly all confounding has been addressed, and thus uncontrolled confounding will remain. The causal estimate is no longer identifiable even after the use of g-computation methods if assumptions including the absence of other biases (such as selection and information bias) are not met. The use of g-computation in this situation will yield a biased effect estimate and may lead to a large-biased effect estimate due to the added influence of other biases. Therefore, considering the sufficient assumptions for g-methods, both acknowledgment and sensitivity analyses to see if and how strongly the assumptions hold, especially within a causal inference framework.

**1.3 Bias Analysis**

Observational studies are subject to random and systematic errors where their very presence could impact the validity and interpretation of results. Random error is minimized and addressed as a result of large sample sizes; yet, addressing systematic error is often only limited to a qualitative discussion section instead of any quantitative bias analysis[47]. As a result, several bias analysis methods primed to address systematic error have been developed, studied, and applied[48-50]. These quantitative bias analysis methods are vital in addressing the uncertainty of empirical results. Since the introduction of core textbooks and chapters describing bias analysis, its application has gradually increased. However, despite the availability of statistical packages, published codes, and online calculators, the application of quantitative bias analysis still remains infrequent[49].

The body of literature on bias analysis is vast, ranging from simple fixed sensitivity analysis to bounding analysis including the recent development of the E-value[51]. Bias analysis methods are generally grouped into internal or external adjustments with the goal of addressing the bias(es) present in the study[52]. External adjustment involves bias formulas (i.e., the quantification of bias from omitting the uncontrolled confounder) to adjust the biased effect estimate. One feature of external adjustment is the ability to apply the bias formula to already summarized results, and thus it does not require the source data with individual-level records. Alternatively, internal adjustment is usually applied to (individual) record-level data, where a bias offset representing the bias factor can be applied to each observation resulting in a bias-adjusted effect estimate after regression[52,53]. Moreover, bias analysis can be understood as a missing data problem, where key information on a confounder is missing, or uncontrolled. One can tangibly derive the unknown variable from a specific probability distribution of other variables to weight the final regression and obtain a bias-adjusted effect estimate.

To employ these methods, information is required to ascertain the bias parameters, or the parameters representing the associations between the uncontrolled confounder, exposure, outcome, and covariates[49,52]. The difficulty in implementing bias analysis hinges on the identification of parameters, rather than the actual analysis itself. Common sources of parameters include literature reviews or expertise given their feasibility in attainment, although internal validation data are most desirable. Bias analysis could be performed using a range of bias parameter values or with distributions of bias parameter values to account for the uncertainty in parameters[54]. Added complexity to bias analysis arises when evaluating multiple sources of biases, or even multiple uncontrolled confounders. In addition to specifying a larger number of parameters for each bias, the order that each bias occurs should also be considered[49,52,55]. Due to

these challenges, many applications of bias analysis have been restricted to the use of a single bias factor applied to a baseline study[39].

There is limited information and application of bias analysis to deal with uncontrolled confounding observed in a time-varying setting with causal inference methods[55-60]. A recent systematic review did not find an example of bias analysis for a time-varying confounder and only a handful of studies evaluated a time-varying exposure[49]. Other tutorials have demonstrated bias analysis with causal inference methods, namely inverse probability weighting, but restricted the bias analysis method to a single point measurement of an uncontrolled confounder[61-63]. Of the existing literature for a time-varying uncontrolled confounder in causal modeling, a proposed method has approached the issue as a missing data problem and hinge on the availability of validation data[57]. The most prominent method involves a bias parameter to augment the IPTW in the fitting of MSM, similar to that of a bias offset[56]. Despite these available methods addressing this issue, their real-world application in practice has been sparse.

The use of sensitivity analysis is critical, especially in the context of causal inference where the strong assumptions are also often untestable. Bias analysis is a transparent way to quantify bias for an effect estimate. Moreover, it gives quantitative guidance on the alternative explanations that could be driving the direction, magnitude, and uncertainty of the empirical results due to systematic errors. This is especially important in the setting of big data in which results may have misleadingly narrow confidence intervals and encourage an overconfidence in results. In a broader public health perspective, bias analysis can help inform areas of future study as well as guide the policies, resource allocation and regulatory decision making that are dependent on valid results[6]. For these reasons, bias analysis remains a critical part of

observational research and necessitates regular application along with further study into gaps of its literature.

### 1.4 Dissertation Structure and Specific Aims

The overarching goals of this dissertation are two-fold (Figure 1.1). First, the methodological goal of this dissertation is to delineate time-varying uncontrolled confounding, which remains understudied despite its importance in the interpretation of causal estimates from complex longitudinal data settings. This provides the rationale for the latter aims to develop, evaluate, and demonstrate two bias analysis methods with g-computation in the fitting of marginal structural models. As there have been limited applications of bias analysis within causal modeling, this dissertation seeks to describe and provide the methods to make bias analysis more accessible to future research. Thus, the final goal of the dissertation includes the public health implications of evaluating the relationships of elevated depressive symptoms and the outcomes of cancer and CVD in a cohort of older adults. Using the methods developed in this dissertation, these relationships will be tested for theorized sources of time-varying uncontrolled confounding to add clarity to the literature.

The specific aim of each study is listed below:

**Study 1**

To describe the causal structures for and quantify the direction and magnitude of time-varying uncontrolled confounding when evaluating causal estimates in time-varying exposure and confounding in cancer epidemiologic studies.

**Study 2**

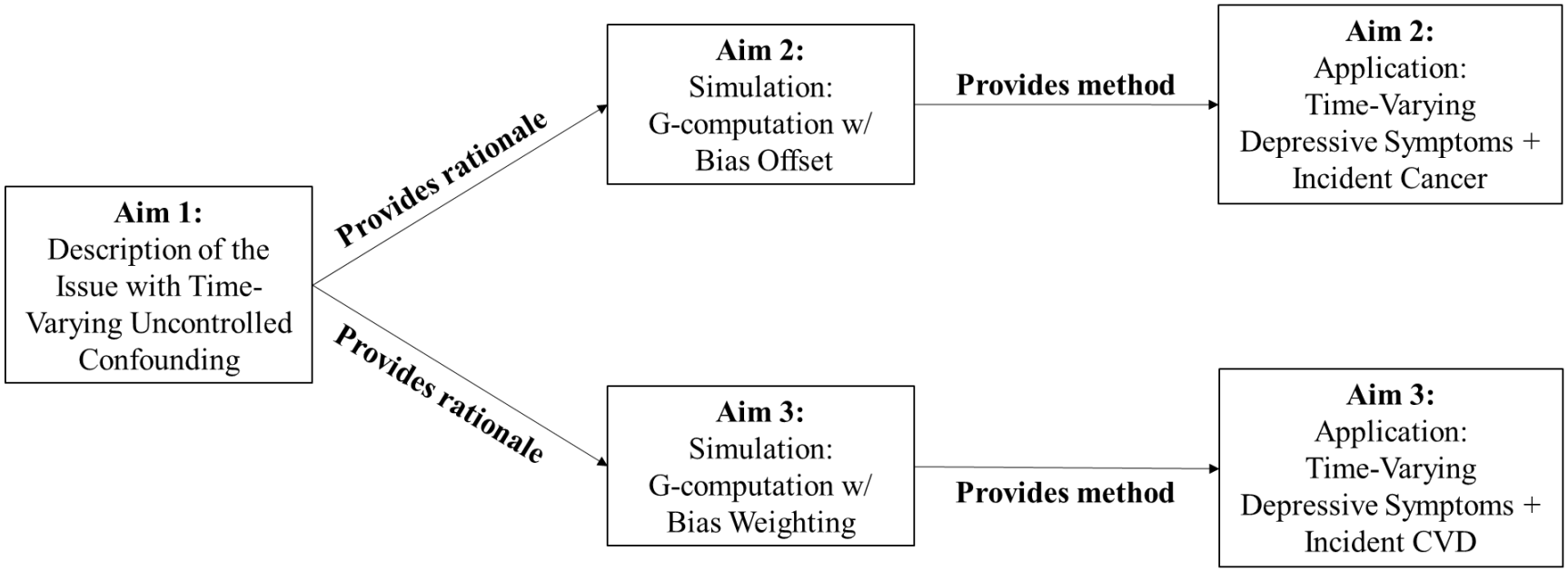
To develop, evaluate and use a novel g-computation algorithm augmented with a bias offset function to adjust for time-varying uncontrolled confounding in simulated datasets and in an

empirical analysis of the effect of elevated depressive symptoms on cancer incidence, including breast and prostate cancers.

**Study 3**

To develop, evaluate and use a novel g-computation algorithm augmented with inverse probability of uncontrolled confounder weighting to adjust for time-varying uncontrolled confounding in simulated datasets and in an empirical analysis of the effect of elevated depressive symptoms on cardiovascular disease incidence.

**Figure 1.1. Relationship of the Specific Aims for this Dissertation**



## 2. Study 1: Time-Varying Uncontrolled Confounding in Cohort Data Analyzed Using G-Computation: Simulated Applications to Cancer Epidemiology

## 2.1 Abstract

**Background:** Uncontrolled confounding is a bias and threat to validity. When using longitudinal time-varying methods including fitting of marginal structural models with g-methods, the required assumption of conditional exchangeability or no uncontrolled confounding extends beyond a baseline (time-fixed) setting and encompasses time-varying study designs. Here, we describe time-varying confounding structures using directed acyclic graphs (DAGs) and illustrate the differences in joint effect estimates obtained from g-computation when the assumption of no time-varying uncontrolled confounding is violated using simulated examples from cancer epidemiology.

**Methods:** We used DAGs to depict and reason about time-varying uncontrolled confounding in cohort study settings. We then used simulations and g-computation to estimate the joint effect of time-varying exposures on the mean difference (MD) and odds ratio (OR) scales. Simulation parameters were sourced from examples in cancer epidemiology. Models including time-varying uncontrolled confounders were considered as the true (bias-adjusted) estimate. Biased joint effect estimates were evaluated by excluding time-varying uncontrolled confounders from models, representing an unmet 'no uncontrolled confounding' assumption. True and biased estimates were compared across a range of simulated magnitudes of uncontrolled confounding to evaluate the impact of time-varying uncontrolled confounding on effect estimates.

**Results:** Time-varying uncontrolled confounding can be described with several scenarios including relationships into subsequently measured confounders, exposures, and the outcome. Estimates were most biased when the uncontrolled confounders were directly related to the outcome. In all simulated illustrations, true and biased estimates differed even when uncontrolled confounding was weak. The greatest differences in joint effects were observed in simulations

where time-varying uncontrolled confounding was consistently strong over several time points. In the illustration of elevated body mass index and obesity-related cancers, the true OR was about1.53, while the most biased estimate was OR [95%SI, simulation interval]: 2.30 [1.91, 2.72] when uncontrolled confounding was strong over all time points.

**Conclusions:** Time-varying uncontrolled confounding has the potential to greatly impact observed joint effect estimates even when using causal inference methods. Given the importance of longitudinal observational studies in advising public health and guidelines, the impact of sources of uncontrolled confounding warrants more recognition and careful evaluation using quantitative bias analysis.

## 2.2 Background

Seminal works have developed the methods for modeling serial measurements of time-varying exposures, outcomes, and confounders when evaluating complex longitudinal and time-varying observational studies[44,64]. In these studies, current measurements are affected by prior measurements and the use of conventional regression can inadvertently bias results[45]. The collection of Robins' generalized methods (g-methods) aptly addresses these complexities through the fitting of marginal structural models to estimate time-varying effects[45]. As a result, there has been recent proliferation in the application of these methods in epidemiology[46].

Confounding remains a threat to validity as it is when an extraneous factor distorts the observed exposure-outcome relationship from the counterfactual[43]. Irrespective of time-fixed or time-varying measurements, confounding requires attention to obtain an unbiased estimate. This is feasible when data are available yet poses challenges when data on the confounder are unavailable. In observational studies, including those with causal inference methods, this uncontrolled confounding is assumed as absent. Prior guides have described the collection of strong and mostly untestable assumptions required for causal inference[45,65]. The no uncontrolled confounding assumption can also be described as conditional exchangeability, where conditional on the covariates sufficient to control for confounding, the exposure is independent of the potential outcome, and those who were exposed would have, on average, the same potential outcome as those who were unexposed[45]. Within time-varying studies, the assumption is expanded and defined as given the prior history of exposure and covariates, the potential outcome is independent of the present exposure[44,45].

The use of causal inference and methods, robust evaluation, and longitudinal data make critical contributions to public health. These studies have the ability to evaluate long-term

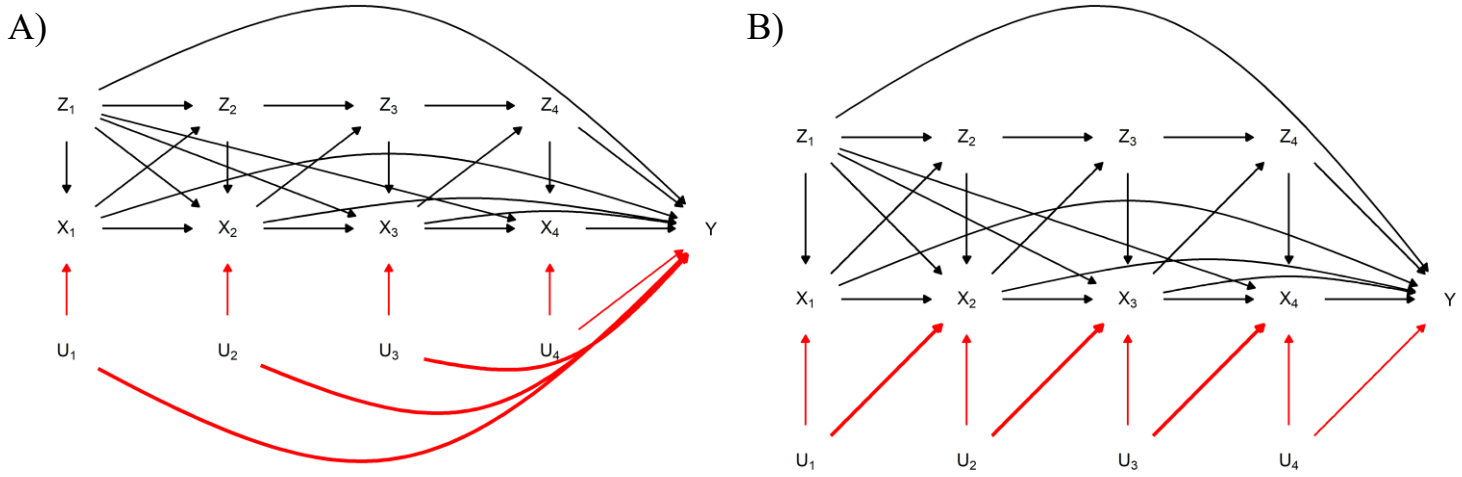
exposures and could potentially inform guidelines as their causal interpretation may be as valued as that of a clinical trial[45]. Yet, uncontrolled confounding is often mentioned in a study's discussion section yet not always formally evaluated. Calls for routine robust evaluation with bias analysis has continued to increase[47,49]. Moreover, the uncontrolled confounding assumption extends beyond a time-fixed measure and may be time-varying in nature. As a result, even if causal methods are used, unmet assumptions including the presence of uncontrolled confounding may bias results, interpretations, and value to public health policy. Thus, this study seeks to describe structures of time-varying uncontrolled confounding and then demonstrate the magnitude of time-varying uncontrolled confounding in simulated data from cancer epidemiology. These illustrations evaluate (i) the mean difference in low-density lipoprotein cholesterol (LDL) with statin non-adherence before and during breast cancer treatment and (ii) the odds of obesity-related cancers with elevated body mass index over childhood to adulthood. This study highlights both time-varying uncontrolled confounding and sensitivity of estimates under different biasing scenarios during using causal modeling in the absence of conditional exchangeability.

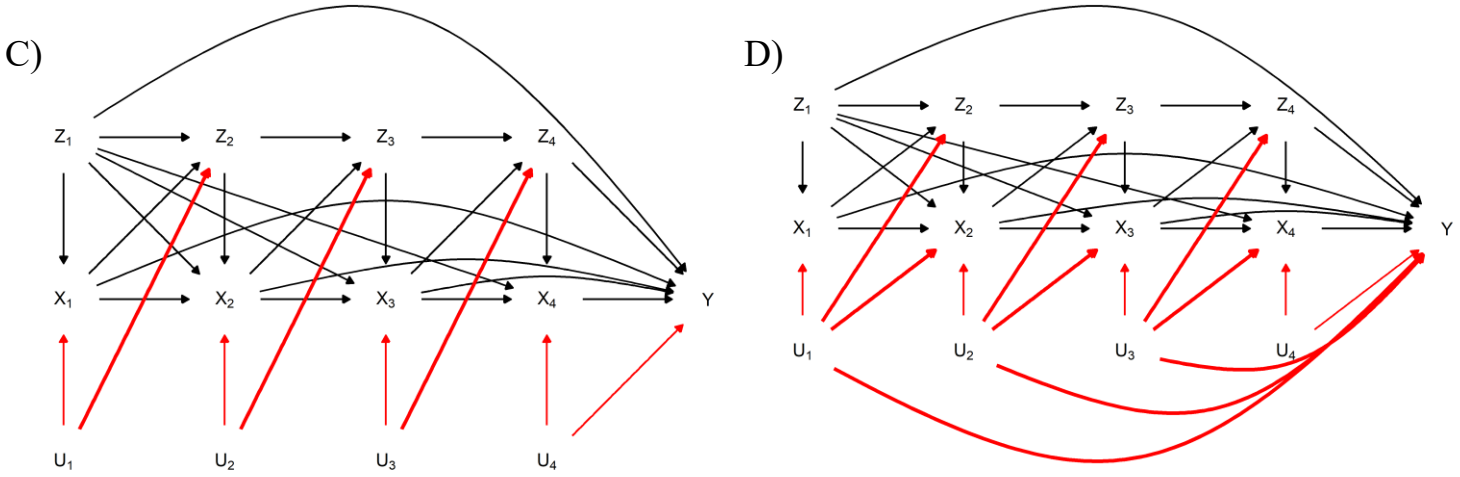
### 2.2.1 Time-Varying Uncontrolled Confounding Structures

The causal relationships between the confounder, exposure, and outcome have been well-described with directed acyclic graphs (DAGs)[64]. These relationships become more complex with multiple time points of measurement. The definition of a time-varying confounder can no longer be limited to an open backdoor path from the exposure and directly into the outcome. A time-varying confounder may also repeatedly affect the exposure and the next exposure measurement or next confounder measurement (Figures 2.1 A-D). In these DAGs, we feature the relationships between our time-varying exposures ($X_t$), measured set of confounders ($Z_t$),

outcome (Y) and uncontrolled confounders ($U_t$) over multiple time points, indexed by time t (1-4). Figure 2.1A shows a classic representation of confounding, where the uncontrolled confounder at time t is associated with the current exposure measurement at time t and the outcome, repeatedly over time. There is no relationship between the uncontrolled confounder at time t with any other exposure or measured confounder in the future. Figure 2.1B represents the relationship where the uncontrolled confounders only associate with the current and subsequent exposure measurements, thereby creating a type of bias often called "M-bias", or a "bow-tie" structure between multiple exposures[43]. Figure 2.1C shows the relationship of the uncontrolled confounders with the current exposure and the next measured confounder. Figure 2.1D represents the combination of Figures 2.1A-C. These figures reflect the time-varying confounding structures suited for causal methodology and estimation of joint, pooled and cumulative treatment effects[45]. Additional variations of these structures could include dependencies between uncontrolled confounders as well as mediating effects through the subsequent uncontrolled confounder (Appendix Figure 2.1).

**Figure 2.1. Directed Acyclic Graphs Depicting the Data-Generating Mechanism and Types of Uncontrolled Confounding of the Joint Effect of $X_t$ and Y over Four Time Points**






Abbreviations: $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 4; $Z_t$: covariates at time t.

## 2.3 Methods

### Notation and Data Generation for Simulations 1.1 and 1.2

The data-generating processes for simulations 1.1 and 1.2 are depicted in the corresponding DAGs in Appendix Figure 2.1 and Figure 2.1A, and Tables 2.1A and B, respectively. Let the subscript *t* represent the time point in which variables were measured, where t = 1 represents the first and baseline value. Let X represent the exposure or treatment, Y as the outcome, Z as the covariate set sufficient for known and measured confounding, and U as the unmeasured and uncontrolled confounder. In our simulations, $Z_1$ represented a collective set of both time-fixed and time-varying confounders[45]. Appendix Figure 2.1 represents a longitudinal design of two time points (simulation 1.1), where $U_1$ affects both the next exposure and outcome and $X_1$ affects $U_2$. To simplify the causal structure over a longer period of follow-up, Figure 2.1A represents four time points (simulation 1.2), in which only $U_t$ affects the current exposure $X_t$ and the outcome.

Using a sample size of N = 50,000, we generated independent variables drawn with a specified probability for each participant. Each variable was iteratively drawn with a specified probability representing the given relationships between previously simulated variables. Unless otherwise noted, variables were drawn under a Bernoulli distribution for binary variables. For all simulations, $U_t$ variables were also included so that the generated data reflected the true causal structure including the U-dependent relationships. All parameters representing known variables were plausibly selected given the illustration and maintain clinically meaningful relationships (Tables 2.1A and 2.1B)[66-68]. The data-generating processes included simulations over a total of K = 1000 replications under a specific seed for reproducibility[69].

**Table 2.1. Parameters for A) Simulation 1.1 for Analyses of Two Time Points and B) Simulation 1.2 for Analyses of Four Time Points**

**A)**

| Variable | Description | Parameters for Simulation 1.1 |
|---|---|---|
| $\mathbf{Z_1}$ | Covariate set before diagnosis (t = 1) | 0.5 |
| $\mathbf{U_1}$ | Uncontrolled confounder before diagnosis (t = 1) | 0.25 |
| $\mathbf{X_1}$ | Exposure: statin non-adherence before diagnosis (t = 1) | $0.2 + 0.1Z_1 + 0.2U_1$ |
| $\mathbf{U_2}$ | Uncontrolled confounder during cancer treatment (t = 2) | $0.25 + 0.2U_1 + 0 \times X_1$ |
| $\mathbf{Z_2}$ | Covariate set during cancer treatment (t = 2) | $0.3 + 0.1Z_1 + 0.15X_1$ |
| $\mathbf{X_2}$ | Exposure: statin non-adherence during cancer treatment (t = 2) | $0.22 + 0.15X_1 + 0.1Z_1 + 0.1Z_2 + 0.2U_1 + 0.2U_2$ |
| **Y** | LDL after cancer treatment | $112 + 10X_1 + 12X_2 + 4X_1X_2 + 2Z_1 + 4Z_2 + \delta_{U1}U_1 + \delta_{U2}U_2$; SD = 25 |

**B)**

| Variable | Description | Parameters for Simulation 1.2 |
|---|---|---|
| $\mathbf{Z_1}$ | Covariate set at childhood (t = 1) | 0.35 |
| $\mathbf{U_1}$ | Uncontrolled confounder at childhood (t = 1) | 0.25 |
| $\mathbf{U_2}$ | Uncontrolled confounder at early adolescence (t = 2) | 0.25 |
| $\mathbf{U_3}$ | Uncontrolled confounder at adolescence (t = 3) | 0.25 |
| $\mathbf{U_4}$ | Uncontrolled confounder at adulthood (t = 4) | 0.25 |
| $\mathbf{X_1}$ | Exposure: overweight at childhood (t = 1) | $\text{expit}(\log(0.10) + \log(1.20)Z_1 + \log(\alpha_{U1})U_1)$ |
| $\mathbf{Z_2}$ | Covariate set at early adolescence (t = 2) | $\text{expit}(\log(0.20) + \log(1.20)Z_1 + \log(1.25)X_1)$ |
| $\mathbf{X_2}$ | Exposure: overweight at early adolescence (t = 2) | $\text{expit}(\log(0.13) + \log(1.25)X_1 + \log(1.20)Z_1 + \log(1.20)Z_2 + \log(\alpha_{U2})U_2)$ |
| $\mathbf{Z_3}$ | Covariate set at adolescence (t = 3) | $\text{expit}(\log(0.20) + \log(1.20)Z_2 + \log(1.25)X_2)$ |
| $\mathbf{X_3}$ | Exposure: overweight at adolescence (t = 3) | $\text{expit}(\log(0.20) + \log(1.25)X_2 + \log(1.20)Z_1 + \log(1.20)Z_3 + \log(\alpha_{U3})U_3)$ |
| $\mathbf{Z_4}$ | Covariate set at adulthood (t = 4) | $\text{expit}(\log(0.20) + \log(1.20)Z_3 + \log(1.25)X_3)$ |

| | | |
|---|---|---|
| **$X_4$** | Exposure: overweight at adulthood (t = 4) | expit(log(0.25) + log(1.25)$X_3$ + log(1.20)$Z_1$ + log(1.20)$Z_4$ + log($\alpha_{U4}$)$U_4$) |
| **Y** | Outcome: Obesity-related cancers | expit(log(0.025) + log(0.93)$X_1$ + log(1.03)$X_2$ + log(1.20)$X_3$ + log(1.33)$X_4$ + log(1.10)$Z_1$ + log(1.10)$Z_4$ + ($\alpha_{U1}$)$U_1$ + ($\alpha_{U2}$)$U_2$ + ($\alpha_{U3}$)$U_3$ + ($\alpha_{U4}$)$U_4$) |

Abbreviations: expit: inverse logit; LDL: low-density lipoprotein cholesterol; SD: standard deviation; t: time; $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome; $Z_t$: covariates at time t.

Footnote: Values for $\delta_{Ut}$ and $\alpha_{Ut}$ are listed in Tables 2.2 and Appendix Table 2.2.

For simulation 1.1, variables were simulated on a linear scale. Y was drawn under a truncated normal distribution with a standard deviation of 25 mg/dL. A lower bound in the data generation process was used to maintain clinically plausible values of LDL. The model for Y is depicted below:

$$Y \sim N(\mu, 25); \mu = 112 + \delta_{X1}X_1 + \delta_{X2}X_2 + \delta_{X1X2}X_1X_2 + \delta_{Z1}Z_1 + \delta_{Z2}Z_2 + \delta_{U1}U_1 + \delta_{U2}U_2$$

For simulation 1.2, the relationships between variables were simulated on the inverse logit scale (i.e., expit). The probability of Y = 1 was generated as:

$$P(Y = 1|X_1, X_2, X_3, X_4, Z_1, Z_4, U_1, U_2, U_3, U_4) =$$

$$expit(\alpha_0 + \alpha_{X1}X_1 + \alpha_{X2}X_2 + \alpha_{X3}X_3 + \alpha_{X4}X_4 + \alpha_{Z1}Z_1 + \alpha_{Z4}Z_4 + \alpha_{U1}U_1 + \alpha_{U2}U_2 + \alpha_{U3}U_3 + \alpha_{U4}U_4)$$

where each $\alpha_V$ is a logit.

To further illustrate a variety of structures and impact of time-varying uncontrolled confounding, simulations akin to that of simulation 1.2 were modified to represent Figures 2.1A-D. The input parameters were the same as Table 2.1B.

**G-computation**

Guides have previously described g-computation methods to estimate the effect estimate and 95% simulation interval (SI)[45,65]. We focused on g-computation without simulation given that the input datasets were simulated and to avoid high-computation issues. We assumed that the assumptions necessary for causal inference including positivity, consistency and no other sources of bias were met in our simulated data. As to examine the presence of uncontrolled confounding, this assumption did not hold. We performed g-computation to obtain the joint direct effects of $X_t$ and Y, i.e., the effect of being exposed at all time points and not through

subsequent exposures, compared with no exposure at all time points. After simulation, we implemented g-computation with the following steps for each replication:

Step 1: Obtain the regression coefficients using regression with an appropriate link. Fit models of all descendants of $X_t$ (i.e., $Z_{t+1}$) with appropriate variables, as well as a model of Y on all $X_t$ and appropriate $Z_t$ and $U_t$ as covariates as defined by the DAGs. For our true (bias-adjusted) illustration, we included $U_t$ to represent the true structure dictated by the DAG as if this data were available.

Step 2: Calculate the potential outcomes under treatment (all $X_t = 1$) and no treatment (all $X_t = 0$) for each subject using the Step 1 parameters.

Step 3: Obtain the causal effect estimate by taking the appropriate contrast between the two potential outcomes on the chosen effect measure scale. The mean of the estimate represents the effect between $X_t$ and Y for each replication. Among all K replications, the median and $2.5^{th}$ and $97.5^{th}$ percentiles of estimates were summarized to represent the causal estimate and 95% simulation interval.

To highlight the consequences of accounting or not accounting for time-varying uncontrolled confounding, we repeated analyses with modifications with respect to $U_t$. $U_t$ were removed in the Step 1 regression and thus the coefficients represented the situation where $U_t$ were conceptually present but remained uncontrolled. Subsequently, Step 2 potential outcomes were calculated in the absence of $U_t$. Step 3 was carried out as described and the results represent the biased estimates. To further highlight the effect of $U_t$, we varied the $U_t$ parameters and repeated simulations. In simulations of four time points, the $U_t$ parameters were of equal magnitude to the exposure and next variable (exposure, confounder, outcome, as applicable). We created a standard scenario called ‘Model 1’ to represent a baseline biasing scenario for each

dataset. Based on the initial $U_t$ inputs, we individually or collectively increased or decreased $U_t$ by a specified magnitude (50%) and repeated analyses. For simulation 1.2, an incremental increase or decrease of 25% over time was also included in biasing scenarios. For all scenarios, increasing or decreasing the $U_t$ parameters represented varying degrees of strong and weak time-varying UC, respectively.

### 2.3.1 Hypothetical Illustrations and Motivation in Cancer Epidemiology

Simulation 1.1 is a hypothetical illustration evaluating the joint effect of statin non-adherence over time and the subsequent difference in LDL among breast cancer patients. As breast cancer patients are often older women who may be burdened with comorbid cardiovascular disease and diabetes, a statin may be recommended to lower LDL given a certain risk profile[66]. The growing interest and research in cardio-oncology has suggested that used of statins may be beneficial on breast cancer prognosis given its antineoplastic properties[70,71]. While statin non-adherence may be observed in the general population, specifically among women with breast cancer, the occurrence of a stressful life event combined with cancer treatment may also result in statin non-adherence[72]. Thus, t = 1 represented statin non-adherence shortly before breast cancer diagnosis, and t = 2 represented statin non-adherence during the breast cancer treatment period. The outcome, Y represented the mean LDL up to two years after finishing treatment. In this example, $Z_t$ embodied the sufficient set to control for confounding, including comorbidities and polypharmacy which may influence statin non-adherence and LDL[73-75]. Parameters were informed by a study that had detailed records on lipids and prescriptions before, during, and after cancer treatment in a cohort of breast cancer patients already on a statin[66]. Furthermore, we chose $U_t$ to represent a risk factor of statin discontinuation and LDL such as

clinical indicators and biomarkers of a statin related adverse event. In this situation, statin use would affect $U_{t+1}$, thereby demonstrating a time-varying confounding structure[74].

Finally, simulation 1.2 is a hypothetical illustration that evaluated the joint odds ratio (OR) of elevated body mass index (overweight or obese status) with obesity-related cancers such as cancers of the colon, rectum, liver, gallbladder, pancreas, and others. This longitudinal follow-up over four time points represented the life course from childhood to adulthood. Several longitudinal cohort studies that combined historical school records with national electronic medical records to examine changes in body weight status during and across specific time periods informed this illustration[67,68,76]. As cancer is suspected to have a long latency period, past studies have hypothesized that the age period in being overweight or obese has an influence on cancer. In this study, $U_t$ represented time-varying uncontrolled confounding, such as markers of poor nutrition and diet which may be unmeasured especially if utilizing electronic medical records or claims data[77]. Given the rare disease assumption, the odds ratio may be approximated as a risk ratio. Thus, we made this assumption when using the risk ratio estimates from literature to inform our parameters to estimate the odds ratio.

Based on the simulated examples, the true joint estimates for simulations 1.1 and 1.2 were approximately a mean difference (MD) of 26.6 mg/dL and an OR of 1.53, respectively. All analyses and figures were performed in R version 4.0.0 and R Studio version 1.2.5042. Given the nature of the study, no data were missing.

## 2.4 Results

### Description of multiple time-varying confounding scenarios

Figures 2.1A-D represent structures of time-varying uncontrolled confounding. Under simulation, we obtained a joint OR [95%SI] of 1.53 [1.23, 1.88] as the true joint estimate of

being always exposed compared to never exposed at all time points. The most biased estimates were observed in Figures 2.1A and 2.1D, where each scenario highlighted the presence of a $U_t$ → Y relationship (Appendix Table 2.1). The combination of all scenarios produced the most biased estimate of joint OR [95%SI]: 2.08 [1.72, 2.53]. Estimates obtained from Figures 2.1B and 2.1C produced mildly biased estimates, each at a OR of approximately 1.60.

**Simulation 1.1**

In our example, we observed a true joint MD [95%SI] in LDL of 26.6 [25.9, 27.2] mg/dL when the patient was statin non-adherent both before and during breast cancer treatment (Appendix Figure 2.2). However, in a plausible case of time-varying uncontrolled confounding, the biased joint effect [95%SI] was 35.3 [34.6, 36.0] mg/dL when the parameter and confounder strengths of $U_1$ and $U_2$ were 16 mg/dL each (Table 2.2). In all models, changing the strength of $U_1$ and $U_2$ demonstrated biased estimates away from the true value, with differences in magnitudes ranging from 4 to 13 mg/dL. Namely, when both $U_1$ and $U_2$ strengths were considerably strong, we observed the most biased MD [95%SI] of 39.7 [38.9, 40.5] mg/dL (Table 2.2). Conversely, when the strengths of $U_t$ were decreased to weaker levels, the resulting estimates were closer to the true MD [95%SI]: 30.9 [30.3, 31.6] mg/dL (Table 2.2), albeit overestimated. Individually examining the regression parameters for each $X_1$ or $X_2$ alone showed similar results, where the $X_2$ in biased models deviated further away from the true estimate.

**Table 2.2. Regression Parameters and the Joint Mean Difference in LDL from Statin Non-Adherence Before and During Breast Cancer Treatment in Simulation 1.1**

| Model | Description | $\delta_{U1}$ | $\delta_{U2}$ | $\delta_{X1}$: $X_1$ [95%SI] | $\delta_{X2}$: $X_2$ [95%SI] | $\delta_{X1X2}$: $X_1X_2$[95%SI] | Joint MD [95%SI] |
|---|---|---|---|---|---|---|---|
| 1 | True | 16 | 16 | 10.6 [9.8,11.3] | 12.0 [11.4,12.5] | 4.0 [3.1,5.0] | 26.6 [25.9,27.2] |
| 1 | Biased | 16 | 16 | 11.8 [11.0,12.5] | 17.8 [17.3,18.4] | 5.7 [4.7,6.7] | 35.3 [34.6,36.0] |
| 2 | $\delta_{U1}$ +50% | 24 | 16 | 12.6 [11.8,13.4] | 19.1 [18.6,19.8] | 6.3 [5.3,7.4] | 38.0 [37.3,38.8] |
| 3 | $\delta_{U2}$ +50% | 16 | 24 | 11.6 [10.8,12.4] | 19.5 [18.9,20.1] | 5.9 [4.8,7.0] | 37.0 [36.3,37.7] |
| 4 | $\delta_{U1}$ & $\delta_{U2}$ +50% | 24 | 24 | 12.4 [11.6,13.2] | 20.7 [20.2,21.4] | 6.5 [5.4,7.6] | 39.7 [38.9,40.5] |
| 5 | $\delta_{U1}$ -50% | 8 | 16 | 11.0 [10.3,11.7] | 16.5 [16.0,17.1] | 5.1 [4.1,6.1] | 32.6 [32.0,33.3] |
| 6 | $\delta_{U2}$ -50% | 16 | 8 | 12.0 [11.2,12.8] | 16.2 [15.7,16.8] | 5.4 [4.4,6.4] | 33.7 [33.0,34.4] |
| 7 | $\delta_{U1}$ & $\delta_{U2}$ -50% | 8 | 8 | 11.2 [10.4,11.9] | 14.9 [14.4,15.5] | 4.8 [3.8,5.8] | 30.9 [30.3,31.6] |

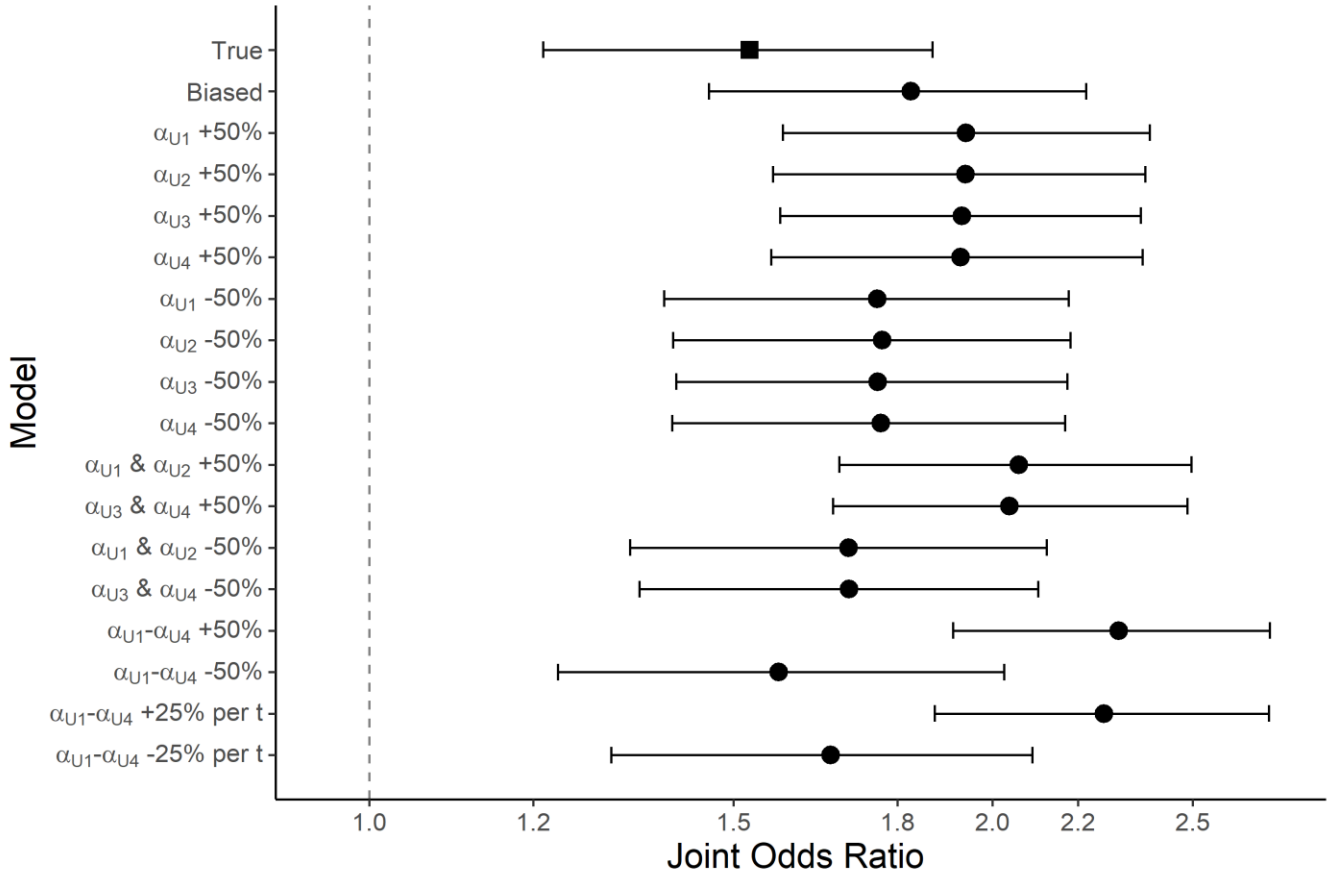
Abbreviations: LDL: low-density lipoprotein cholesterol; MD: Mean Difference; SI: simulation interval; $X_t$: statin non-adherence at time t.

Footnote: $\delta_{Ut}$ represents the strength of relationship between $U_t$ and the respective variable on the corresponding association measure scale. In all analyses, statin adherence was considered as the referent group

**Simulation 1.2**

We expanded our study design to evaluate being overweight or obese from childhood to adulthood on the odds ratio of obesity-related cancer incidence. We observed that being overweight/obese across all time points compared to being of normal weight across all time points, resulted in a true bias-adjusted joint OR [95%SI] of 1.53 [1.21, 1.87]. Yet, leaving $U_t$ uncontrolled resulted in a joint OR [95%SI] of 1.83 [1.46, 2.22] in model 1 (Figure 2.2, Appendix Table 2.2). For models 2-5, where $U_t$ was individually modified at a single time point to a strong confounder, we observed biased estimates of 1.93. For models 6-9, $U_t$ was deemed weak at individual time points and yielded biased results. The largest differences from the true value were observed in the presence of multiple strong confounders. For example, models 10 and 11 had strong confounding early (i.e., t = 1-2) or late (i.e., t = 3-4), respectively and both resulted in biased joint estimates (OR [95%SI]: 2.06 [1.69, 2.50] and 2.04 [1.68, 2.48], respectively). Strong confounding over all time points and gradually increasing until the end of follow-up yielded the largest biased estimate (OR [95%SI]: 2.30 [1.91, 2.72] and 2.26 [1.88, 2.72], respectively). Finally, the presence of weaker uncontrolled confounding over all time points produced closer results to that of the true estimate (OR [95%SI]: 1.58 [1.23, 2.03]).

**Figure 2.2. Simulated Illustration of the Joint Odds Ratio of Overweight and Obesity Status with Related Cancers over Follow-up Across Multiple Biasing Scenarios in Simulation 1.2**




Footnote: $\alpha_{Ut}$ represents the strength of relationship between $U_t$ and the respective variable on the corresponding association measure scale. In all analyses, being non-overweight or non-obese was considered as the referent group

## 2.5 Discussion

We examined the structures and impact of time-varying uncontrolled confounding that violate the conditional exchangeability condition in causal inference. Time-varying confounding, including uncontrolled confounding, can present under various causal structures and may result in differing degrees of biased results. Structures with direct relationships between the uncontrolled confounders and outcome gave the most biased results. In the simulated illustrations, the differences in magnitude between the biased estimates (expectation of the biased estimator) and the true values were exacerbated in the presence of strong uncontrolled confounding. We observed that models with a greater number of time points with multiple strong uncontrolled confounders yielded biased estimates. In the presence of weaker time-varying uncontrolled confounding, we observed smaller deviations in the biased estimate. These hypothetical examples reflected studies in cancer epidemiology and support the importance of understanding and addressing time-varying uncontrolled confounding in a context-dependent manner.

Studies using g-methods assume no uncontrolled confounding. Authors may speculate about the strength and impact of potential uncontrolled confounding, yet the use of quantitative sensitivity analyses remains infrequent[52]. Only recently has there been a greater call for more bias analysis methods in the field[49]. We demonstrated that time-varying uncontrolled confounding may present itself in several structures beyond the typical exposure-confounder-outcome relationship. Thus, authors should theorize sources, direction, and relationship of uncontrolled confounding within the context of the study given that different structures have differing magnitudes of biased effects as demonstrated in the present study. Yet, in all causal analyses, a detailed understanding about the system and observed data is recommended as to best

evaluate and target causal questions[78]. In our simulated examples, we theorized sources of uncontrolled confounding within the context of the study and used DAGs to illustrate the relationships of uncontrolled variables on each association. Often, such causal analyses have been featured in big data such as medical claims, registries, or electronic medical records that have strengths in sample size and length of follow-up[46]. These sources were not built for research purposes and may fall short in critical laboratory, social or lifestyle data, thus leaving the potential for uncontrolled confounding in analyses. In this era of big data, databases are revered for their covariate-rich data and size yet as they may inform subsequent policy allocations. It is best to fully understand and test results including the robustness to different biasing scenarios, as to not lead to a potentially misguided or overconfident conclusion. For these reasons, sensitivity analyses are especially important in studies including causal inference methods.

In simulation 1.1, we examined changes in LDL from statin non-adherence that occurred before and during breast cancer treatment in simulated data. We showed that excluding time-varying uncontrolled confounding resulted in deviations from the true estimate. In this illustration, statin non-adherence in multiple time points resulted in a greater increase in LDL for breast cancer patients, which was worsened by strong time-varying uncontrolled confounding. Prior to simulations, we indicated that clinical biomarkers of adverse events may be unmeasured given a specific data source and pose as time-varying uncontrolled confounders influencing both non-adherence and LDL. These biased estimates shed light on the impact of uncontrolled confounding and understanding of statin non-adherence and LDL in this population[66]. In the general population, elevated LDL levels are associated with a higher risk of cardiovascular events and remain prominent contributors to clinical risk scores used to guide treatment, including statin intensification[79,80]. Women with breast cancer, may also be at risk of subsequent

cardiovascular events given the hypothesized cardiotoxic effects of cancer treatment, as well as better prognosis and survival observed with this cancer[70]. Therefore, careful examination of treatments and adherence in older women, including those with breast cancer and changes in LDL remain critical. These longitudinal relationships should be examined considering time-varying uncontrolled confounding as to better understand the intersection of cancer and cardiovascular disease epidemiology.

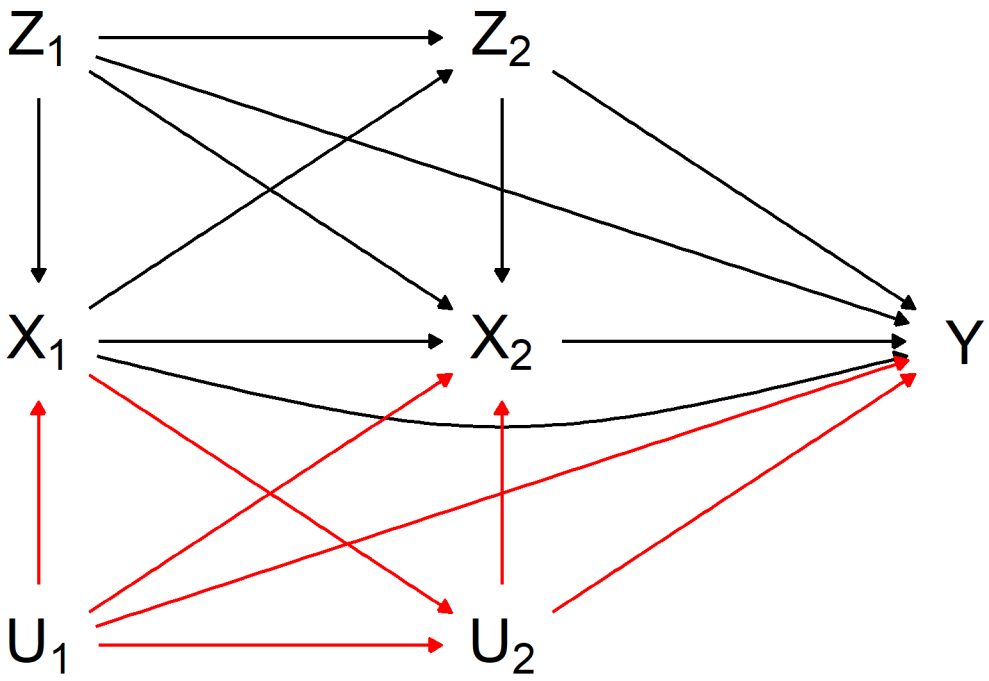
For simulation 1.2, studies have investigated elevated body mass index over one's life course, often finding strong relationships of chronic overweight/obese status during certain periods of one's life with specific cancers[67,68,76]. In the context of the global obesity epidemic and long latency of cancer, it remains of interest to study these relationships over long periods. Like simulation 1.1, strong time-varying uncontrolled confounding resulted in the biased effect estimates. The most biased ORs were observed when there were multiple time-varying uncontrolled confounders. Within the context of this hypothetical study and plausibility in the field, there may only be a few confounders that would have as strong of a relationship with obesity and cancer that are consistently present over a life course. In this example, we suggested that time-varying uncontrolled confounding represented markers of poor nutrition and diet associated with both obesity and cancer. Persistence of overweight or obese status from adolescence into adulthood has been documented and it may be suggested that poorer lifestyle habits developed in youth carry forward into adulthood[81,82]. Like the prior studies, birth cohorts and data registries may not have this type of dietary data readily available and thus remain subject to uncontrolled confounding. This example highlights potential uncontrolled confounding by poor nutrition and further support identifying policy related actions to address overweight and obese status including during youth.

Our study strengths include the evaluation of different measures of association across varying strengths of confounding. For our illustrations, we utilized literature-informed inputs to suggest what may be observed in the presence of time-varying uncontrolled confounding. Finally, simulations allowed for the systematic evaluation of the impact of time-varying uncontrolled confounding. Our study also has some limitations. In applying g-computation, we assumed that other causal conditions than uncontrolled confounding held in our analyses. Given that our data were simulated, we were able to dictate and address some assumptions in the data-generating process. Practically, we acknowledge that assumptions may be untestable and would require sensitivity analyses. We assumed that our covariate $Z_t$, represented a sufficient set of confounders in our examples, however we note it is a combination of variables that individually may have varying relationships with other variables in the system. The use of a single variable may not fully reflect the magnitude and direction of confounding present in observational data.

Ultimately, uncontrolled confounding is a critical causal concern that warrants sensitivity analysis. We observed that the presence of multiple time-varying uncontrolled confounders resulted in a larger biased estimate in our simulated models. We highlighted that time-varying uncontrolled confounding can present in differing structures and magnitudes given measured variables. Given the increasing use of advanced causal models in large datasets to study many public health problems that involve interventions that change over time, rigorous quantitative bias analysis of time-varying uncontrolled confounding in such settings is indispensable.

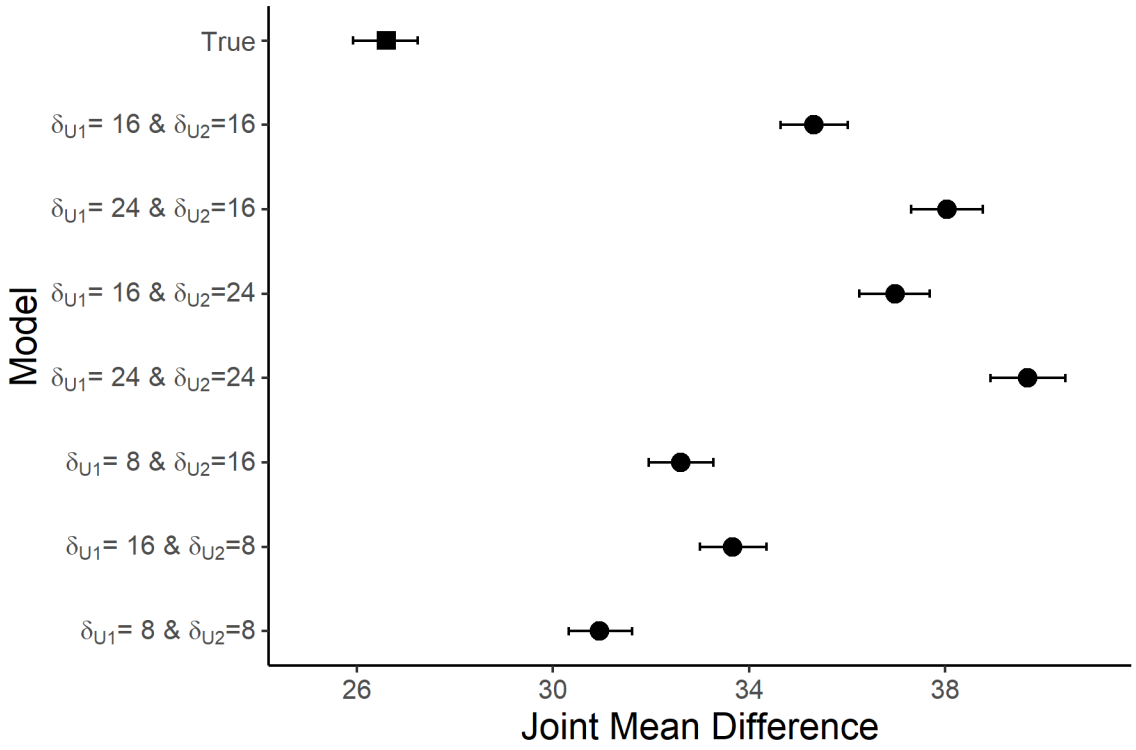
## 2.6 Appendix

**Appendix Figure 2.1. Directed Acyclic Graph Depicting the Data-Generating Mechanism and Uncontrolled Confounding of the Joint Association of $X_t$ and Y over Two Time Points**




Abbreviations: $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 2; $Z_t$: covariates at time t

**Appendix Figure 2.2. Simulated Illustration of the Joint Mean Difference in LDL from Statin Non-Adherence Before and During Breast Cancer Treatment in Simulation 1.1**




Footnote: $\delta_{Ut}$ represents the strength of relationship between $U_t$ and the respective variable on the corresponding association measure scale. In all analyses, statin adherence was considered as the referent group

**Appendix Table 2.1. Joint Odds Ratio Over Follow-up Across Multiple Confounding Scenarios**

| Model | Confounding Description | Estimate | Joint OR [95%SI] |
|---|---|---|---|
| | | True | 1.53 [1.23,1.88] |
| A | $X_t \leftarrow U_t \rightarrow Y$ | Biased | 1.83 [1.48,2.26] |
| B | $X_t \leftarrow U_t \rightarrow X_{t+1}$ | Biased | 1.59 [1.19,2.04] |
| C | $X_t \leftarrow U_t \rightarrow Z_{t+1}$ | Biased | 1.60 [1.20,2.07] |
| D | Combination of A, B and C | Biased | 2.08 [1.72,2.53] |

Abbreviations: OR: Odds Ratio; SI: simulation interval; $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 4; $Z_t$: covariates at time t

Footnote: $\alpha_{Ut} = \log(1.60)$ for all models and represents the strength of relationship between $U_t$ and the respective variable. Confounding descriptions apply to all time points except for t = 4.

**Appendix Table 2.2. Joint Odds Ratio of Overweight and Obesity Status with Related Cancers over Follow-up Across Multiple Biasing Scenarios**

| Model | Description | $\alpha_{U1}$ | $\alpha_{U2}$ | $\alpha_{U3}$ | $\alpha_{U4}$ | Joint OR [95%SI] |
|---|---|---|---|---|---|---|
| 1 | True | log(1.60) | log(1.60) | log(1.60) | log(1.60) | 1.53 [1.21,1.87] |
| 1 | Biased | log(1.60) | log(1.60) | log(1.60) | log(1.60) | 1.83 [1.46,2.22] |
| 2 | $\alpha_{U1}$ +50% | log(2.02) | log(1.60) | log(1.60) | log(1.60) | 1.94 [1.58,2.38] |
| 3 | $\alpha_{U2}$ +50% | log(1.60) | log(2.02) | log(1.60) | log(1.60) | 1.94 [1.57,2.37] |
| 4 | $\alpha_{U3}$ +50% | log(1.60) | log(1.60) | log(2.02) | log(1.60) | 1.93 [1.58,2.36] |
| 5 | $\alpha_{U4}$ +50% | log(1.60) | log(1.60) | log(1.60) | log(2.02) | 1.93 [1.56,2.36] |
| 6 | $\alpha_{U1}$ -50% | log(1.26) | log(1.60) | log(1.60) | log(1.60) | 1.76 [1.39,2.18] |
| 7 | $\alpha_{U2}$ -50% | log(1.60) | log(1.26) | log(1.60) | log(1.60) | 1.77 [1.40,2.18] |
| 8 | $\alpha_{U3}$ -50% | log(1.60) | log(1.60) | log(1.26) | log(1.60) | 1.76 [1.41,2.17] |
| 9 | $\alpha_{U4}$ -50% | log(1.60) | log(1.60) | log(1.60) | log(1.26) | 1.77 [1.40,2.17] |
| 10 | $\alpha_{U1}$ & $\alpha_{U2}$ +50% | log(2.02) | log(2.02) | log(1.60) | log(1.60) | 2.06 [1.69,2.50] |
| 11 | $\alpha_{U3}$ & $\alpha_{U4}$ +50% | log(1.60) | log(1.60) | log(2.02) | log(2.02) | 2.04 [1.68,2.48] |
| 12 | $\alpha_{U1}$ & $\alpha_{U2}$ -50% | log(1.26) | log(1.26) | log(1.60) | log(1.60) | 1.70 [1.34,2.13] |
| 13 | $\alpha_{U3}$ & $\alpha_{U4}$ -50% | log(1.60) | log(1.60) | log(1.26) | log(1.26) | 1.71 [1.35,2.10] |
| 14 | $\alpha_{U1}$-$\alpha_{U4}$ +50% | log(2.02) | log(2.02) | log(2.02) | log(2.02) | 2.30 [1.91,2.72] |
| 15 | $\alpha_{U1}$-$\alpha_{U4}$ -50% | log(1.26) | log(1.26) | log(1.26) | log(1.26) | 1.58 [1.23,2.03] |
| 16 | $\alpha_{U1}$-$\alpha_{U4}$ +25% per time t | log(1.60) | log(1.80) | log(2.08) | log(2.50) | 2.26 [1.88,2.72] |
| 17 | $\alpha_{U1}$-$\alpha_{U4}$ -25% per time t | log(1.60) | log(1.42) | log(1.30) | log(1.22) | 1.67 [1.31,2.09] |

Abbreviations: OR: Odds Ratio; SI: simulation interval

Footnote: $\alpha_{Ut}$ represents the strength of relationship between $U_t$ and the respective variable. In all analyses, being non-overweight or non-obese was considered as the referent group

## 3. Study 2: Investigating Whether Elevated Depressive Symptoms Affect Cancer Outcomes: G-Computation with Bias Analysis to Offset Time-Varying Uncontrolled Confounding

### 3.1 Abstract

**Background:** Quantitative bias analysis of the impact of uncontrolled confounding is typically used for single time point studies, if at all. The prior studies of depressive symptoms and cancer outcomes are mixed, of a single time point, and subject to residual or uncontrolled confounding. We extend the bias offset framework to a time-varying setting combined with g-computation using simulations and investigating cancer outcomes.

**Methods:** We simulated 50,000 individuals with repeated measures to evaluate the impact of time-varying uncontrolled confounding ($U_t$) on the joint risk difference (RD). We used g-computation including and excluding $U_t$ to fit marginal structural models. We demonstrated the use of a bias offset in combination with g-computation to recover the true joint estimate. In an illustration, we investigated the association between elevated depressive symptoms and incident cancer (overall, breast, and prostate) in a cohort of 12,502 older adults from the Health and Retirement Study in 1998 and 2000. We fit regression models with demographics and time-varying lifestyle factors for our g-computation. We hypothesized time-varying sedentary behavior as the $U_t$. Finally, using plasmode simulations, we performed quantitative bias analysis offsetting varying bias factors for uncontrolled sedentary behavior.

**Results:** In our simulations, time-varying uncontrolled confounding yielded more biased estimates when the $U_t$ were excluded from the model. The use of the bias offset was successful in removing the bias due to $U_t$ and recovered the original estimate. In our illustration, elevated depressive symptoms in both time points resulted in a joint RD [95%SI, simulation interval] of -0.005 [-0.049, 0.046]. We observed robust results with the bias offset analyses when using stronger magnitudes of bias factors. Likewise, similar results were observed in gender-specific analyses for breast and prostate cancers.

**Conclusions:** This study demonstrated the use of a bias offset in combination with g-computation to evaluate time-varying uncontrolled confounding in simulated datasets. Our illustration suggested that the results were robust to time-varying uncontrolled confounding after using the bias offset. The expansion of this bias offset method to time-varying exposure and confounding settings merits future investigation and application in causal inference studies.

## 3.2 Introduction

The era of big data has encouraged evaluation of intricate research questions especially in observational studies with repeated measurements. These study designs include complex relationships, where covariates are affected by a prior exposure measurement and also act as confounders for the current exposure measurement[45]. Thus, g-methods such as g-computation are important for estimation of the effects of time-varying exposures. In g-computation, several assumptions are required. One assumption, conditional exchangeability or no uncontrolled confounding, becomes more complex when evaluating time-varying study designs. Not only is it assumed that there is no uncontrolled confounding of the current exposure measurement, but this assumption is expanded such that the definition encompasses no uncontrolled confounding given the history of prior variables[45].

The conditional exchangeability assumption is particularly strong and untestable in observational study settings. Quantitative methods have been developed to assess the impact of biases[43,47]. When evaluating single time point, or baseline studies, bias analysis methods can evaluate the impact of uncontrolled confounding, measurement error, and selection bias on the observed results[47,52]. One popular method involves a bias offset. A bias factor is defined as the amount of bias due to a hypothesized source (i.e., uncontrolled confounding) and is subtracted (or offset) from the final regression internally on a record-level[53,83]. Recently, there have been calls from the research community to increase the use of sensitivity analysis methods as a more routine or standard practice in research. These sensitivity methods to rigorously subject results to sources of bias remain imperative to research, and only become more important in the context of big data and causal inference methods in time-varying studies.

To date, there are few methods and applications that evaluate biases in combination with causal inference methods and time-varying studies[56]. In this study, we extend the bias analysis framework of the bias offset to a time-varying setting to evaluate the presence of time-varying uncontrolled confounding. Through plasmode simulation combined with g-computation, we apply these methods to several illustrations evaluating the relationship of elevated depressive symptoms and incident cancer from a large cohort of older adults.

### 3.3 Methods

#### Notation and definitions

The parameters for simulations 2.1-2.3 are depicted in Table 3.1 and reflect the data-generating mechanisms of Figures 3.1A-C. We generated a sample of N = 50,000 individuals where each binary variable was drawn under a Bernoulli distribution. Data generation was repeated over K = 1000 replications.

We used the following notation in our directed acyclic graph (DAGs) for each t, time point: let X represent the exposure, Y represent the outcome, Z represent the covariate set sufficient for known and measured confounding, and U represent the uncontrolled confounder. For all simulations, we examined two time points. The DAGs represent the following scenarios: one uncontrolled confounder (Figure 3.1A, simulation 2.1), two time-varying uncontrolled confounders (Figure 3.1B, simulation 2.2) and two time-varying uncontrolled confounders where $U_2$ also acts as a mediator for $X_1$ (Figure 3.1C, simulation 2.3). To highlight the bias due to uncontrolled confounding and maintain a consistent joint effect across all analyses, simulations 2.1 and 2.3 were simulated with a null effect including all direct and indirect pathways. Simulation 2.3 was also simulated with non-null indirect effects as to further support the bias analysis model, however the joint effect thus varied.

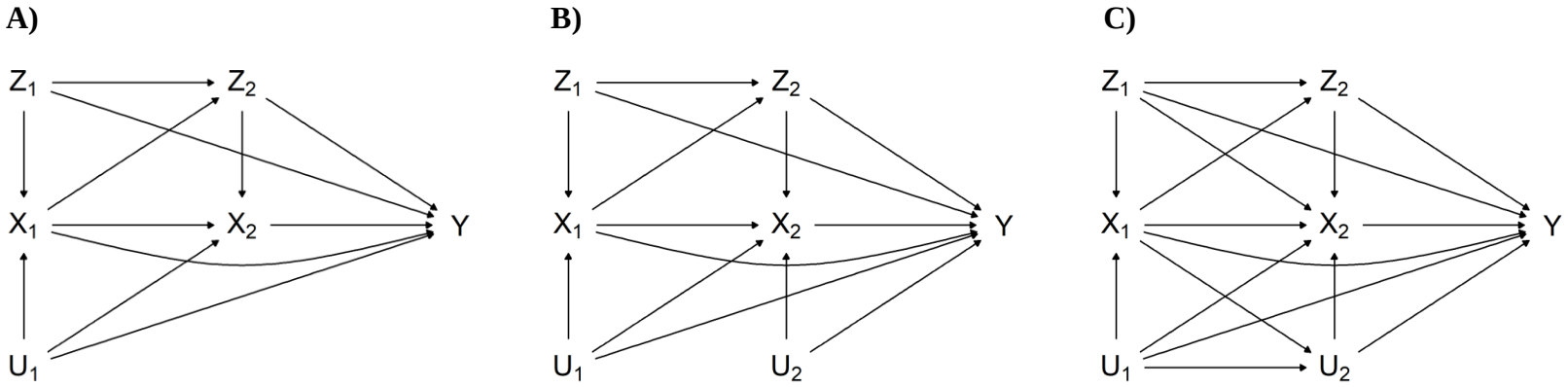
**Table 3.1. Parameters for Simulations 2.1-2.3**

| Variable | Description | Simulation 2.1:<br><br>Time-varying exposures with 1 uncontrolled confounder | Simulation 2.2:<br><br>Time-varying exposures with 2 uncontrolled confounders | Simulation 2.3:<br><br>Time-varying exposures with 2 uncontrolled confounders and a mediator |
|---|---|---|---|---|
| $U_1$ | Uncontrolled confounder at time = 1 | 0.4 | 0.6 | 0.4 |
| $Z_1$ | Covariate at time = 1 | 0.5 | 0.4 | 0.5 |
| $X_1$ | Exposure at time = 1 | $0.4 + 0.1Z_1 + R_{U1}U_1$ | $0.4 + 0.2Z_1 + R_{U1}U_1$ | $0.4 + 0.1Z_1 + R_{U1}U_1$ |
| $U_2$ | Uncontrolled confounder at time = 2 | | 0.6 | $0.2 + 0.2U_1 + R_{X1}X_1$ |
| $Z_2$ | Covariate at time = 2 | $0.2 + 0.1Z_1 + 0\times X_1$ | $0.25 + 0.2Z_1 + 0.3X_1$ | $0.2 + 0.1Z_1 + R_{X1}X_1$ |
| $X_2$ | Exposure at time = 2 | $0.15 + 0.2X_1 + 0.1Z_2 + R_{U1}U_1$ | $0.1 + 0.25X_1 + 0.2Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $0.15 + 0.2X_1 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ |
| Y | Outcome after time = 2 | $0.1 + 0\times X_1 + 0\times X_2 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1$ | $0.05 + 0.08X_1 + 0.08X_2 + 0.04X_1X_2 + 0.15Z_1 + 0.15Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $0.1 + 0\times X_1 + 0\times X_2 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ |

Abbreviations: $R_{X1:}$ parameter linking $X_1$ and the specified variables; $R_{Ut}$: parameter linking $U_t$ and the specified variables; $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 2; $Z_t$: covariates at time t

Footnote: To demonstrate a null joint effect with simulation 2.1 and 2.3 across all bias analysis models, each indirect relationship (i.e., $X_1 \rightarrow Z_2$ and $X_1 \rightarrow U_2$) was also set to null. Secondary analyses of simulation 2.3 added $R_{X1} = 0.1$ for each indirect relationship, whereby the joint effect changed with each bias analysis model. All variables were generated as binary variables.

**Figure 3.1. Directed Acyclic Graphs of Time-Varying Exposures Depicting A) Only One Uncontrolled Confounder, B) Only Two Time-Varying Uncontrolled Confounders and C) Two Time-Varying Uncontrolled Confounders and an Unmeasured Mediator**




Abbreviations: $U_t$; uncontrolled confounder at time t, $X_t$; exposure at time t, Y; outcome after time = 2, $Z_t$; covariates at time t

**G-computation**

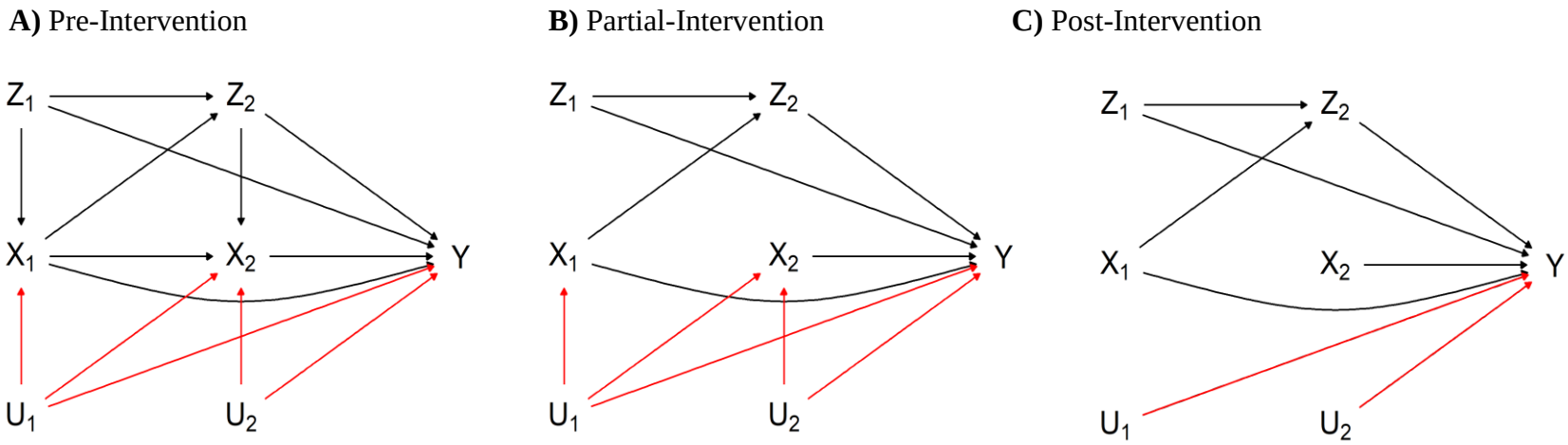
Methods for g-computation have been previously described[45,65]. Given the simulated nature of the data, we assumed that all assumptions held, except for conditional exchangeability. In summary, we fit regressions models for the descendants of X, and our outcome using the appropriate set of variables. We calculated the potential outcomes under hypothetical treatment ($X_t = 1$) and no treatment ($X_t = 0$) at each time point. To obtain the joint risk difference (RD), we contrasted the potential outcomes when treated at both time points vs never treated at any time point. This joint estimate represented the effect of exposure on the outcome, not through the subsequent exposure. We took the mean value from each replication, and then summarized the distribution of 1000 replications. The median, $2.5^{th}$ and $97.5^{th}$ percentiles of the 1000 replications represented our final estimate and 95% simulation interval (SI) on the risk difference scale. Bias and root mean square error (RMSE) for the joint estimates were calculated from the estimates of each replication. As to contrast the obtained effect estimates with respect to uncontrolled confounding, $U_t$ were included in relevant regression modeling to represent our 'True' estimate and $U_t$ were excluded from relevant regression modeling to represent our 'Biased' estimate. The latter estimate represents the case when the regression models are mis-specified and in violation of the conditional exchangeability assumption. To examine varying effects of uncontrolled confounding, we individually varied the simulation parameters relating to $U_t$.

**Bias Analysis with a Bias Offset**

We extend the use of a bias offset to a time-varying setting with g-computation. Figure 3.2A (or Figure 3.1B) represents the pre-intervention DAG and the data generating process for the simulation. In g-computation, the post-intervention DAG reflects the randomization of each exposure at each time point (Figure 3.2C). In the presence of time-varying uncontrolled

confounding, we propose an intermediate step termed a partial-intervention DAG, where each exposure is intervened on for only known relationships (Figure 3.2B). Therefore, the uncontrolled confounding paths are preserved at each time point and recognized in the relationships. Using the bias offset, the relationships of the time-varying uncontrolled confounders are subtracted, or offset from the observed outcome and result in the post-intervention DAG. To achieve Figures 3.2A-C, we used simulation to develop a plasmode dataset with the g-computation framework to estimate the causal effect estimate, offset by time-varying uncontrolled confounding.

**Figure 3.2. Directed Acyclic Graphs Depicting A) Pre-Intervention, B) Partial-Intervention after use of the Bias Offset and C) Post-Intervention after G-Computation**




Abbreviations: $U_t$; uncontrolled confounder at time t, $X_t$; exposure at time t, Y; outcome after time = 2, $Z_t$; covariates at time t

The following steps describe the plasmode simulation:

Step 1: From the prior g-computation steps, obtain the regression parameters representing the relationships for the descendants of X and outcome modeled from the biased regression models (i.e., 'Biased' model that excluded $U_t$). Let $\delta_{Ut}$ represent the regression parameters relating Y to $U_t$, conditional on all other exposures and covariates (i.e., obtained from the 'True' regression model). Finally, regress $U_t$ on all exposures and covariates as: $U_1 \sim X_1+X_2+Z_1+Z_2$ and $U_2 \sim X_1+X_2+Z_1+Z_2+U_1$, on the specified (e.g., linear) scale.

Step 2: Generate the plasmode dataset through simulating the exposure and the time-varying covariates. The exposures are simulated according to their marginal probability, thereby representing the post-intervention DAG. Each $U_t$ is simulated according to their marginal probability (i.e., a variable called sim_$U_t$). Any descendant of the exposure (i.e., $Z_2$) will be simulated using the simulated exposures and the regression parameters from Step 1.

Step 3: Calculate the probability of each $U_t$ with the regression parameters from Step 1 and the newly simulated exposures and covariates.

Step 4: Calculate the probability of Y with the biased regression parameters from Step 1 and the newly simulated exposures and covariates.

Step 5: Calculate the bias offsets for each $U_t$ as follows: $\delta_{Ut} \times [P(U_t|.) - \text{sim\_}U_t]$. This represents the bias due to $U_t$. This will also maintain the original marginal probability of Y.

Step 6: Subtract the bias offset for each $U_t$ from the probability of Y. Simulate a U-free Y under a Bernoulli distribution using this new probability.

Step 7: Run the regression for the final marginal structural model for the desired effect estimate using the simulated plasmode data, outcome, and exposures. This is repeated for all

replications. Summarize the distribution of estimates for all replications to obtain the point estimate and 95% simulation interval.

### 3.3.1 Illustration with the Health and Retirement Study

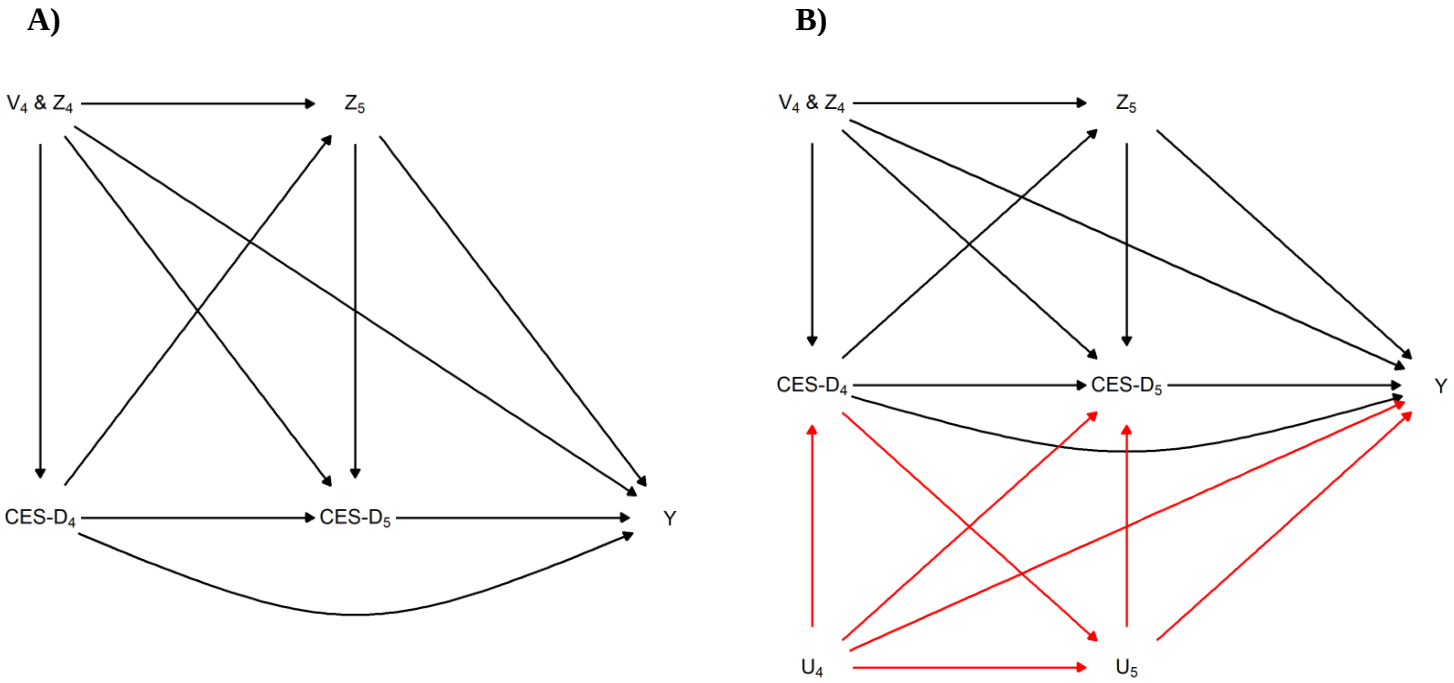
To evaluate our methods, we used data from the Health and Retirement Study (HRS). In this illustration, we evaluated the joint relationship of elevated depressive symptoms in with incidence of overall cancer, breast cancer, and prostate cancer and with g-computation and then applied our bias offset method in evaluating time-varying uncontrolled confounding from sedentary behavior.

The construction of the HRS has been previously described elsewhere[84]. In summary, we utilized the RAND summarized data files (2018 v1) to identify a cohort of patients in 1998 and 2000, the earliest years of data harmonization. The survey waves in 1998 and 2000 correspond to waves 4 and 5, respectively and survey responses in 1998 were considered as baseline in this illustration[85]. Survey data and follow-up were available until 2018. Excluded participants were age ineligible (<50 years old), had partially completed or missing survey data on depressive symptoms, with prior cancer, missing censoring data, and missing key data on demographics and covariates in the 1998 and 2000 survey waves. The final analytical cohort comprised 12,502 older adults (Appendix Figure 3.1).

The Center for Epidemiological Studies Depression scale (CES-D) short form survey were conducted at each wave. A CES-D score ≥ 4 was considered as having elevated depressive symptoms at the time of the survey[86-88]. Incident cancer was defined as an affirmative response to having cancer or malignant tumor (except skin cancer) since the last survey wave (i.e., after 2000). Moreover, we evaluated gender-specific outcomes of breast and prostate cancer among females and males, respectively using the same criteria. Using the HRS Restricted Data,

respondents were asked to identify the primary cancer site[89]. Affirmative responses to either breast or prostate cancer as the primary cancer site during the same wave of reporting incident cancer were used to inform these outcomes. Based on prior literature and theoretical considerations, we used the following baseline and time-varying covariates for confounding: baseline age, gender, race (white/non-white), ethnicity, high-school equivalent education (≥ 12 years vs <12 years) and time-varying married/partnered status, current alcohol status (drinker/non-drinker, defined by the number of drinks per day when drinking), current smoking status (smoker/non-smoker), body mass index and performing vigorous physical activity at least 3 times/week (Figure 3.3A)[11,15,17,90]. Site-specific analyses and covariates were the same as the overall cohort, except for the breast cancer analyses where parity (number of children ever born) was included as a baseline covariate.

**Figure 3.3. Directed Acyclic Graphs for the Health and Retirement Study Assessing the Joint Effect of Elevated Depressive Symptoms and Cancer Incidence A) with and B) without Time-Varying Uncontrolled Confounders**




Abbreviations: CES-D: Center for Epidemiological Studies Depression scale score; $V_4$: baseline covariates at wave 4: age, gender, race, ethnicity, high-school equivalent education (and parity); $Z_4$ and $Z_5$: time-varying covariates at wave 4 and 5, respectively: martial/partnered status, alcohol status, smoking status, body mass index and vigorous physical activity; Y: incident cancer (overall, breast, or prostate) after wave 5; $U_4$ and $U_5$: time-varying uncontrolled sedentary behavior at wave 4 and 5, respectively.

**Analyses with the Health and Retirement Study**

Clinical characteristics were summarized as mean ± standard deviation or frequency (percentage, %), where applicable. Descriptive characteristics were stratified by elevated depressive symptoms (CES-D <4 vs ≥ 4) at baseline wave 4, as well as in wave 4 and 5.

We used g-computation as previously described to obtain the joint risk difference of persistent elevated depressive symptoms with our series of outcomes, compared with never having elevated depressive symptoms. First, we fit logistic regression models for descendants of our baseline exposure and the outcome including all two-way interaction terms across exposures and baseline and time-varying covariates. These parameters along with the marginal probabilities of exposures were used in simulation over 1000 replications. The regression for the final marginal structural model estimated the joint RD, representing the joint risk difference of elevated depressive symptoms in 1998 and 2000 with incident cancer, breast cancer, or prostate cancer as applicable.

In this example, we used bias analysis to evaluate time-varying uncontrolled confounding from sedentary behavior (Figure 3.3B). Using prior literature and information on lack of physical activity to inform the magnitude of parameters, we simulated sedentary behavior into the HRS and performed sensitivity analyses according to the steps described above[20,91-93]. Given that sedentary behavior was unmeasured, we chose bias parameters of 0.005 and 0.0025 for waves 4 and 5, respectively for the development of the U-free Y. These parameters were additionally multiplied by 1.5x and 2x to represent low, moderate, and strong uncontrolled confounding. For breast and prostate cancer analyses, linear regression with a normal link were used given challenges in convergence in the final marginal structural models. The plasmode dataset was replicated 1000 times. To further examine potential model misspecification, we

evaluated several intermediate models for our g-computation where two-way interaction terms were only A) included between the main exposures, B) between the main exposures and baseline covariates, and C) between the main exposure, baseline covariates, and current exposure-current time-varying covariate to supplement our final model of all two-way interactions. Finally, we addressed potential misclassification of the exposure by defining elevated depressive symptoms as CES-D ≥ 3, which has been noted as a more sensitive and commonly used cut-off within the HRS population[88].

Analyses and figures were performed in R version 4.0.0 and R Studio version 1.2.5042, and SAS version 9.4 (Cary, NC). For HRS analyses, complete case analyses were used as only 2.5% of participants were excluded from the eligible population of cancer-free individuals with CES-D data in 1998-2000. The UCLA Institutional Review Board approved this study for access to the HRS Restricted Data.

### 3.4 Results

#### 3.4.1 Simulations Demonstrating the Bias Offset

**Simulation 2.1**

Results comparing the true, biased, and offset bias analysis are presented in Table 3.2 for one uncontrolled confounder. For all analyses, we observed a true joint risk difference of 0.00 as expected. When the strength of $U_1$ was 0.05, the observed joint RD [95%SI] was 0.005 [-0.004, 0.015] and thus the relative bias was small. Incrementally increasing $U_1$ resulted in more biased estimates. When $U_1$ was 0.20, the observed biased RD [95%SI] was 0.072 [0.062, 0.081]. However, for each of the models irrespective of the level of uncontrolled confounding, the final RD [95%SI] after use of the bias offset was approximately 0.000.

**Table 3.2. Simulation Results for the Joint Risk Difference in Simulation 2.1**

| $R_{U1}$ | Model | Model | Joint RD [95%SI] | Bias | RMSE |
|---|---|---|---|---|---|
| 0.05 | 1 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | Biased | 0.005 [-0.004,0.015] | 0.005 | 0.007 |
| | | Bias-offset | 0.001 [-0.014,0.014] | 0.000 | 0.007 |
| 0.1 | 2 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | Biased | 0.019 [0.009,0.028] | 0.019 | 0.020 |
| | | Bias-offset | 0.000 [-0.013,0.014] | 0.000 | 0.007 |
| 0.15 | 3 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | Biased | 0.042 [0.032,0.051] | 0.042 | 0.042 |
| | | Bias-offset | 0.000 [-0.014,0.014] | 0.000 | 0.007 |
| 0.2 | 4 | True | 0.000 [-0.009,0.009] | 0.000 | 0.005 |
| | | Biased | 0.072 [0.062,0.081] | 0.072 | 0.072 |
| | | Bias-offset | 0.000 [-0.014,0.015] | 0.000 | 0.008 |

Abbreviations: RD: risk difference; RMSE: root mean square error; $R_{U1}$: risk difference for the relationship between $U_1$ and the respective variables; SI: simulation interval

Footnote: The true joint effect is set to 0.00 in simulations. True models represent results from correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where U were excluded. Bias-offset models represent the incorrectly specified g-computation model after offsetting bias due to U.

**Simulation 2.2**

Results for the true, biased, and bias-offset models are presented in Table 3.3. Representing simulation 2.2, the true joint RD was approximately 0.245 in simulations. Incrementally increasing both $U_t$ parameters by 0.05, as well as evaluating the presence of one strong and one weak confounder, all resulted in biased estimates. Having two strong confounders resulted in a joint RD [95%SI] of 0.347 [0.337, 0.357], while having one strong then one weak uncontrolled confounder yielded a joint RD [95%SI] of 0.318 [0.308, 0.327]. Yet, use of the bias offset to address the time-varying uncontrolled confounding, the observed joint RD [95%SI] was approximately 0.245 for all models, albeit slightly wider confidence intervals than before.

**Simulation 2.3**

Results comparing the true, biased, and bias-offset models are presented in Table 3.4 for simulation 2.3. Under a simulated null joint effect, we see similar results to prior simulations. The presence of two strong uncontrolled confounders resulted in worse results (joint RD [95%SI]: 0.116 [0.107, 0.126]). After bias analysis with the bias offset, the final estimates were similar to the true joint RD, with an observed relative bias of <0.013 in the most extreme case. In addition, we further examined a non-null indirect effect. Results were like those of the null joint effect, where strong confounders led to more biased results. Use of the bias offset could similarly address the uncontrolled confounding.

**Table 3.3. Simulation Results for the Joint Risk Difference in Simulation 2.2**

| $R_{U1}$ | $R_{U2}$ | Model | Model | Joint RD [95%SI] | Bias | RMSE |
|---|---|---|---|---|---|---|
| 0.05 | 0.05 | 1 | True | 0.245 [0.236,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.252 [0.243,0.261] | 0.007 | 0.008 |
| | | | Bias-offset | 0.245 [0.231,0.261] | 0.000 | 0.008 |
| 0.1 | 0.1 | 2 | True | 0.245 [0.235,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.271 [0.261,0.280] | 0.026 | 0.026 |
| | | | Bias-offset | 0.245 [0.230,0.261] | 0.000 | 0.008 |
| 0.15 | 0.15 | 3 | True | 0.244 [0.236,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.302 [0.293,0.312] | 0.057 | 0.057 |
| | | | Bias-offset | 0.245 [0.229,0.260] | 0.000 | 0.008 |
| 0.2 | 0.2 | 4 | True | 0.245 [0.235,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.347 [0.337,0.357] | 0.102 | 0.102 |
| | | | Bias-offset | 0.241 [0.224,0.257] | -0.004 | 0.009 |
| 0.2 | 0.05 | 5 | True | 0.245 [0.235,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.318 [0.308,0.327] | 0.073 | 0.073 |
| | | | Bias-offset | 0.247 [0.231,0.261] | 0.001 | 0.008 |
| 0.05 | 0.2 | 6 | True | 0.245 [0.236,0.255] | 0.000 | 0.005 |
| | | | Biased | 0.280 [0.271,0.290] | 0.035 | 0.036 |
| | | | Bias-offset | 0.245 [0.230,0.261] | 0.000 | 0.008 |

Abbreviations: RD: risk difference; RMSE: root mean square error; $R_{Ut}$: risk difference for the relationship between $U_t$ and the respective variables; SI: simulation interval

Footnote: The true joint effect is set to 0.245 in simulations. True models represent results from correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where $U_t$ were excluded. Bias-offset models represent the incorrectly specified g-computation model after offsetting bias due to U.

**Table 3.4. Simulation Results for the Joint Risk Differences in Simulation 2.3**

| | | | | Null Joint Effect | | | Non-Null Joint Effect | | |
|---|---|---|---|---|---|---|---|---|---|
| $R_{U1}$ | $R_{U2}$ | Model | Model | Joint RD [95%SI] | Bias | RMSE | Joint RD [95%SI] | Bias | RMSE |
| 0.05 | 0.05 | 1 | True | 0.000 [-0.009,0.009] | 0.000 | 0.004 | 0.015 [0.006,0.024] | 0.000 | 0.005 |
| | | | Biased | 0.008 [-0.001,0.017] | 0.008 | 0.009 | 0.023 [0.014,0.032] | 0.008 | 0.010 |
| | | | Bias-offset | 0.001 [-0.012,0.014] | 0.001 | 0.007 | 0.016 [0.002,0.030] | 0.001 | 0.007 |
| 0.2 | 0.2 | 2 | True | 0.000 [-0.009,0.010] | 0.000 | 0.005 | 0.030 [0.020,0.041] | 0.000 | 0.005 |
| | | | Biased | 0.116 [0.107,0.126] | 0.116 | 0.116 | 0.146 [0.136,0.157] | 0.116 | 0.116 |
| | | | Bias-offset | 0.013 [-0.003,0.028] | 0.013 | 0.015 | 0.043 [0.027,0.058] | 0.013 | 0.015 |
| 0.2 | 0.05 | 3 | True | 0.000 [-0.008,0.009] | 0.000 | 0.005 | 0.015 [0.005,0.025] | 0.000 | 0.005 |
| | | | Biased | 0.078 [0.068,0.087] | 0.077 | 0.078 | 0.092 [0.082,0.102] | 0.077 | 0.077 |
| | | | Bias-offset | 0.004 [-0.010,0.020] | 0.004 | 0.008 | 0.019 [0.004,0.034] | 0.004 | 0.009 |
| 0.05 | 0.2 | 4 | True | 0.000 [-0.009,0.008] | 0.000 | 0.005 | 0.030 [0.021,0.040] | 0.000 | 0.005 |
| | | | Biased | 0.039 [0.030,0.048] | 0.039 | 0.039 | 0.070 [0.061,0.080] | 0.040 | 0.040 |
| | | | Bias-offset | 0.005 [-0.008,0.020] | 0.005 | 0.009 | 0.036 [0.021,0.050] | 0.006 | 0.009 |

Abbreviations: RD: risk difference; RMSE: root mean square error; $R_{Ut}$: risk difference for the relationship between $U_t$ and the respective variables; SI: simulation interval

Footnote: The true joint null effect is set to 0.00 in simulations. In additional analyses, the indirect effect was non-null and thus the true joint estimates would vary in each model. True models represent results from correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where $U_t$ were excluded. Bias-offset models represent the incorrectly specified g-computation model after offsetting bias due to U.

### 3.4.2 Clinical Illustrations with the Health and Retirement Study

The cohort comprised 12,502 participants with a mean±SD age of 65±9 years old, and included 60% females. Cohort descriptives are presented in Table 3.5 and Appendix Table 3.1. Baseline elevated depressive symptoms were observed in 15% of subjects. The elevated depressive symptoms group at baseline included more females, persons of Hispanic ethnicity, not married/partnered, current smokers and participated in less vigorous physical activity. After both survey waves, 7% of subjects had sustained or persistent elevated depressive symptoms, while 78% of subjects never had elevated depressive symptoms in this timeframe. Across both survey waves, persons with sustained elevated depressive symptoms were less likely to participate in vigorous physical activity and the prevalence decreased between waves 4 to 5.

**Table 3.5. Cohort Descriptives of 12,502 Health and Retirement Study Participants**

**Stratified by Baseline Level of Depressive Symptoms**

| Characteristic | Total | No elevated depressive symptoms: CES-D <4 | Elevated depressive symptoms: CES-D ≥4 |
|---|---|---|---|
| N (%) | 12,502 | 10,665 (85.3%) | 1837 (14.7%) |
| Age (years) | 64.8 ± 9.3 | 64.8 ± 9.2 | 65.3 ± 9.8 |
| 50-<60 | 4265 (34.1%) | 3624 (34.0%) | 641 (34.9%) |
| 60-<70 | 4426 (35.4%) | 3812 (35.7%) | 614 (33.4%) |
| 70-<80 | 2865 (22.9%) | 2476 (23.2%) | 389 (21.2%) |
| ≥80 | 946 (7.6%) | 753 (7.1%) | 193 (10.5%) |
| Gender (% Female) | 7455 (59.6%) | 6147 (57.6%) | 1308 (71.2%) |
| Race (%) | | | |
| White | 10416 (83.3%) | 9038 (84.7%) | 1378 (75.0%) |
| African American | 1687 (13.5%) | 1313 (12.3%) | 374 (20.4%) |
| Other | 399 (3.2%) | 314 (2.9%) | 85 (4.6%) |
| Hispanic (%) | 907 (7.3%) | 649 (6.1%) | 258 (14.0%) |
| Married/Partnered (%) | 8657 (69.2%) | 7650 (71.7%) | 1007 (54.8%) |
| Geographic Region (%) | | | |
| Northeast | 2069 (16.6%) | 1733 (16.3%) | 336 (18.3%) |
| Midwest | 3260 (26.1%) | 2907 (27.3%) | 353 (19.2%) |
| South | 4963 (39.7%) | 4130 (38.7%) | 833 (45.3%) |
| West | 2199 (17.6%) | 1885 (17.7%) | 314 (17.1%) |
| Education (years) | 12.3 ± 3.2 | 12.5 ± 3.0 | 10.8 ± 3.6 |
| Education ≥12 years (%) | 9236 (73.9%) | 8197 (76.9%) | 1039 (56.6%) |
| Lifestyle Factors | | | |
| Alcohol use (%) | | | |
| Non-drinker | 8612 (68.9%) | 7159 (67.1%) | 1453 (79.1%) |
| Low/Moderate | 2493 (19.9%) | 2301 (21.6%) | 192 (10.5%) |
| Heavy | 1397 (11.2%) | 1205 (11.3%) | 192 (10.5%) |
| Body Mass Index ($kg/m^2$) | 27.2 ± 5.1 | 27.1 ± 5.0 | 28.0 ± 5.8 |
| <18.5 | 164 (1.3%) | 134 (1.3%) | 30 (1.6%) |
| 18.5-<25 | 4231 (33.8%) | 3683 (34.5%) | 548 (29.8%) |
| 25-<30 | 5044 (40.3%) | 4352 (40.8%) | 692 (37.7%) |
| 30-<35 | 2178 (17.4%) | 1805 (16.9%) | 373 (20.3%) |
| ≥35 | 885 (7.1%) | 691 (6.5%) | 194 (10.6%) |
| Current Smoker (%) | 2013 (16.1%) | 1631 (15.3%) | 382 (20.8%) |
| Physical Activity (%) | 5898 (47.2%) | 5273 (49.4%) | 625 (34.0%) |

Footnote : Data presented as N (%) or mean±standard deviation where applicable

Over the median [IQR, interquartile range] follow-up of 13.8 [7.8, 18.0] years, there were 2,277 incident cancer cases. After accounting for time-varying confounding with g-computation, we observed a joint RD [95%SI] of -0.005 [-0.049, 0.046] (Table 3.6). In sensitivity analyses, we theorized that time-varying sedentary behavior could be an important yet uncontrolled confounder. Using a series of bias offset parameters, we demonstrated that time-varying sedentary behavior would minimally impact our observed estimates. Even theorizing a strong magnitude of bias in the context of sedentary behavior yielded robust results. We similarly evaluated outcomes of breast cancer and prostate cancer among the subcohort of 7442 females and 5047 males, respectively. Over the same follow-up, there were 348 incident breast cancer cases and 488 incident prostate cancer cases. We observed a joint RD [95%SI] of -0.023 [-0.043, 0.001] for breast cancer using g-computation, assuming our assumptions held. For the outcome of prostate cancer, we observed a joint RD [95%SI] of -0.016 [-0.080, 0.077]. Like the main analyses, evaluation for sedentary behavior also did not change the observed estimates. Similar relationships were observed for time-varying uncontrolled confounding for sedentary behavior. For all analyses, estimates were imprecise and had wide confidence intervals.

**Table 3.6. Illustrative Example Investigating the Relationship of Elevated Depressive Symptoms (CES-D≥4) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior**

| | Joint RD [95%SI] | | |
|---|---|---|---|
| Outcome | Overall Cancer | Breast Cancer | Prostate Cancer |
| N Event/N | 2277/12,502 | 348/7442 | 488/5047 |
| | | | |
| Main Analyses | -0.005 [-0.049,0.046] | -0.023 [-0.043,0.001] | -0.016 [-0.080,0.077] |
| Bias Analysis with Bias Offset | | | |
| Low | -0.006 [-0.054,0.040] | -0.020 [-0.041,0.003] | -0.020 [-0.084,0.065] |
| Moderate | -0.006 [-0.051,0.036] | -0.020 [-0.040,0.007] | -0.021 [-0.085,0.061] |
| Strong | -0.008 [-0.054,0.038] | -0.019 [-0.039,0.006] | -0.020 [-0.081,0.064] |

Abbreviations: SI: simulation interval; RD: risk difference; CES-D: Center for Epidemiological Studies Depression scale score

Footnote: Lower depressive symptoms was the reference in all analyses. Fitted logistic regression for g-computation included covariates described in text and additionally included interaction terms between exposures, baseline covariates and time-varying covariates. Site-specific analyses utilized normal link in final regressions.

Bias factors were set to 0.005 and 0.0025 for waves 4 and 5, respectively in the low bias offset model. The moderate and strong bias offset models multiplied each bias factor by 1.5 and 2, respectively.

We further examined the outcome model specification through the inclusion of hierarchical levels of interaction terms (Appendix Table 3.2). Generally, we observed similar conclusions to that of our main analyses. Notably, in our overall cancer analyses, model specification that only included interaction terms for the exposures suggested a positive joint relationship even in quantitative bias analysis, yet all analyses included wide confidence intervals as well as small estimates of magnitude. In additional sensitivity analyses evaluating the lower definition of elevated depressive symptoms (CES-D ≥3) (22% at baseline), the joint relationship of overall cancer and site-specific were akin to that of our main analyses (Appendix Table 3.3). Morevoer, the use of the same bias factors in the series of bias analysis models likewise demonstrated robust results. Testing other model specifications for the evaluation of CES-D ≥3 further suggested a positive relationship with overall cancer when only including exposure interactions (Appendix Table 3.4).

**3.5 Discussion**

In this study, we extended bias offset methods to a time-varying setting using plasmode simulations and g-computation. We illustrated these methods using illustrations from the Health and Retirement Study to investigate the joint relationships of elevated depressive symptoms and incident overall, breast, and prostate cancers and theorized time-varying sedentary behavior as the uncontrolled confounders.

We propose and demonstrated the use of a bias analysis method in combination with causal inference methods to assess time-varying uncontrolled confounding in the violation of the conditional exchangeability assumption. Using these plasmode simulations, we believe that this method could assist researchers in quantitatively assessing bias when using causal inference methods such as g-computation. As research continues to expand into using more advanced

methods, the need for equally strong sensitivity analyses is imperative in testing the robustness of results to sources of bias. We chose to focus on marginal risk difference as the effect estimate of interest given their usefulness in causal inference as well as interpretability. Moreover, we chose to evaluate the risk difference because the linear scale can easily be subtracted (or offset) in this bias analysis. However, given the interest in conditional effect measures, as well as logistic or log-linear effect estimates (i.e., the odds ratio and risk ratio), additional study and testing would be needed to better evaluate the use of this specific bias offset for those estimates.

The evaluation of the bias offset in our clinical example using data from the Health and Retirement Study was hard to investigate. For all analyses, confidence intervals were large and difficult to reach a conclusion. Additional studies are warranted using a larger sample size. Nevertheless, our results are similar to that of Niles and O'Donovan, who also investigated depressive symptoms in the HRS[90]. During their evaluation of depressive symptoms in 2006 and 2008, they did not observe a relationship for cancer incidence. Their analyses did not consider time-varying confounding and investigated a later period of time; however, the concordance of our results highlight this null relationship. Furthermore, meta-analyses and other studies that had investigated baseline or chronic/recurrent depressive symptoms likewise suggested null to small relationships with equally wide confidence intervals[11,15,17,21]. These mixed results were also observed for our site-specific analyses and possibly corroborate other studies that had investigated similar questions[15,22]. Within our g-computation and outcome models, we included baseline demographics and time-varying covariates of lifestyle factors such as smoking, body mass index, alcohol use and physical activity, where the latter factors were often missing from prior studies. The previous studies that demonstrated positive relationships may have been subject to residual confounding given their limited set of confounders in their analyses[10,12].

Finally, we evaluated sedentary behavior as a potential time-varying uncontrolled confounder. Studies have called for improved distinction between definitions of physical activity and sedentary behavior, as they are not direct opposites in meaning. To date, there is little information specifically pertaining to this question in the context of depressive symptoms and cancer incidence[20]. A recent consortium established in 2021 highlighted the need for more research in the relationships of these psychosocial factors such depressive symptoms and health behaviors of sedentary behavior with cancer incidence. With our bias analysis, we used prior studies to estimate the sedentary behavior prevalence in older adults and direction of relationship to cancer[94-96]. Moreover, physical inactivity and sedentary behavior are distinct health behaviors, but we also used the data on physical inactivity to act as a proxy of validation data to estimate and inform the magnitude of parameters in our analyses. Within this bias analysis, it was suggested that the observed results were robust to uncontrolled confounding by time-varying sedentary behavior. Additional studies are needed to first understand the relationship of sedentary behavior, depressive symptoms, and cancer incidence, as well as any time-varying relationship.

These analyses are not without limitations. There are several assumptions in g-computation beyond conditional exchangeability. We assumed that we met these assumptions including correct model specification in all analyses. In our simulations, we were able to correctly specify these models. For our HRS analyses, we chose to use logistic regression for our intermediate models in the g-computation given its strengths in obtaining valid probabilities and included several interaction terms to fit a more flexible model. We further evaluated model misspecification by including fewer interaction terms and observed generally similar results except for the model that only included exposure interaction. Additional methods including

fitting fully saturated regression models or supervised learning techniques may be more useful in model specification. The magnitudes of effects in our HRS analyses were small with wide confidence intervals, where uncontrolled confounding may be less of a concern in this situation. We assumed that we met all assumptions, however other potential confounders critical in cancer epidemiology were limited, including information on diet and family history or genetics. We acknowledge potential overadjustment specifically in site-specific analyses as we chose to utilize multiple variables and potential confounders within the rich HRS data. Given the smaller sample size and exposures groups, studies of a larger scale would be needed to better investigate this research question. Finally, for all analyses, we chose fixed bias parameters and varied values in sensitivity analyses. However, these methods do not consider random error[52]. The use of probabilistic bias analysis methods may be more informative in addressing the uncertainty in chosen bias parameters as well as accounting for random error.

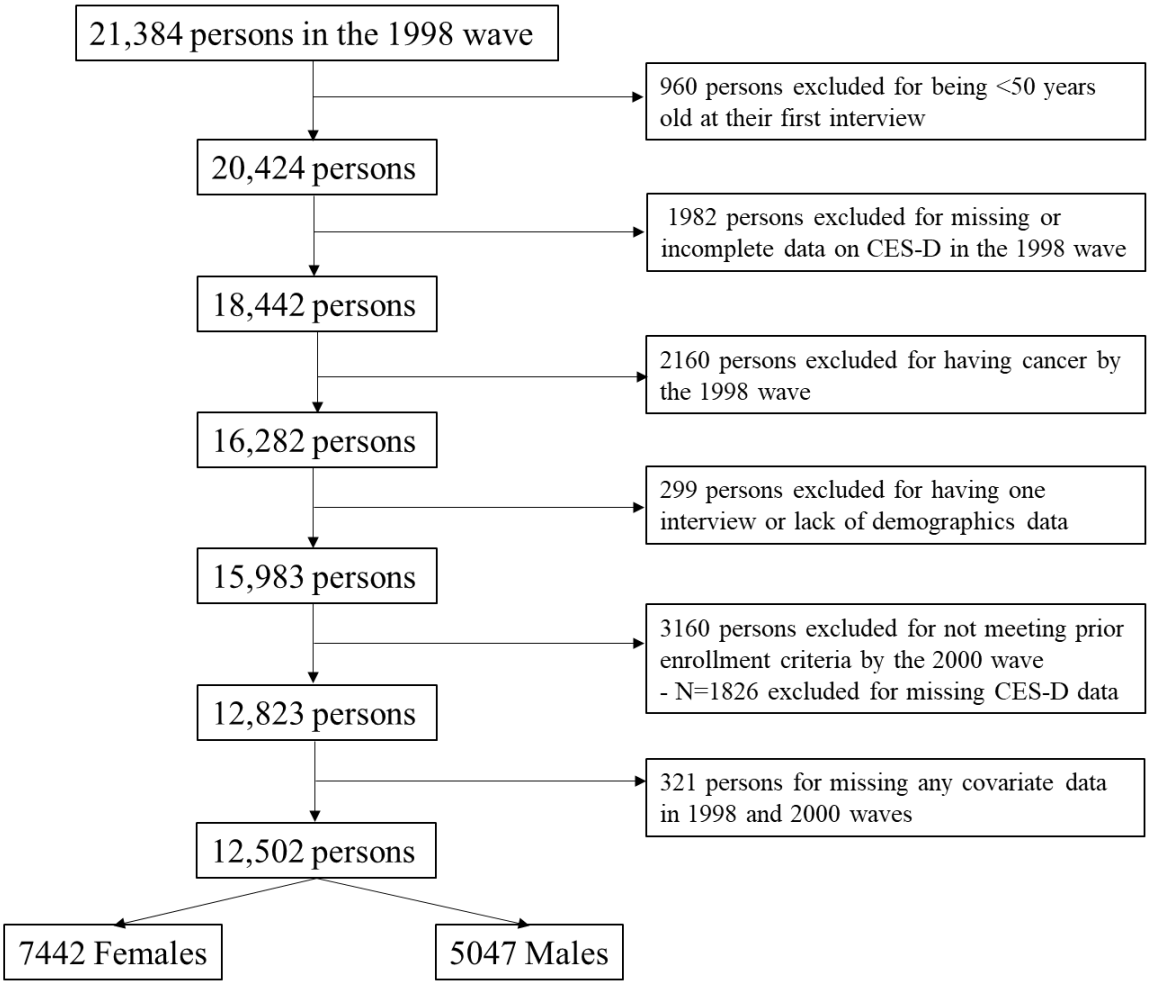
G-computation holds many strengths in the fitting of marginal structural models. The use of outcome modeling, compared with exposure modeling in the case of inverse probability treatment weighting, can be more efficient with more narrow confidence intervals. Moreover, the use of data from the HRS includes strengths of nationally representing the US. In addition, the HRS strengths include having the breadth and depth of granular survey data including repeated measures that afforded the ability to evaluate a time-varying and longitudinal analysis.

Thus, this study demonstrated the use of a bias offset in combination with g-computation to evaluate time-varying uncontrolled confounding in simulated datasets. In clinical application, the bias offset and g-computation were more difficult to evaluate, and thus further studies are warranted. Considering the growth in causal inference methods, we hope that the use of

plasmode simulation and bias analysis encourages researchers to implement this framework and utilize more sensitivity methods.

### 3.6 Appendix

**Appendix Figure 3.1. Cohort Construction for Health and Retirement Study Analyses**




Abbreviations: CES-D: Center for Epidemiological Studies Depression scale score

Footnote: N = 13 females were excluded from the breast cancer analyses due to missing data on parity.

**Appendix Table 3.1. Cohort Descriptives of 12,502 Health and Retirement Study Participants Stratified by Elevated Depressive Symptoms Across Waves 4 and 5**

| Characteristic | No elevated depressive symptoms in either Wave 4 or 5 | Elevated depressive symptoms in Wave 4 only | Elevated depressive symptoms in Wave 5 only | Elevated depressive symptoms in both Wave 4 and 5 |
|---|---|---|---|---|
| N (%) | 9745(78.0%) | 950 (7.6%) | 920 (7.4%) | 887 (7.1%) |
| Age (years) | 64.7 ± 9.1 | 64.7 ± 9.7 | 65.1 ± 10.1 | 65.9 ± 9.8 |
| 50-<60 | 3290 (33.8%) | 352 (37.1%) | 334 (36.3%) | 289 (32.6%) |
| 60-<70 | 3540 (36.3%) | 313 (32.9%) | 272 (29.6%) | 301 (33.9%) |
| 70-<80 | 2253 (23.1%) | 198 (20.8%) | 223 (24.2%) | 191 (21.5%) |
| ≥80 | 662 (6.8%) | 87 (9.2%) | 91 (9.9%) | 106 (12.0%) |
| Gender (% Female) | 5493 (56.4%) | 637 (67.1%) | 654 (71.1%) | 671 (75.6%) |
| Race (%) | | | | |
| White | 8308 (85.3%) | 730 (76.8%) | 730 (79.3%) | 648 (73.1%) |
| African American | 1156 (11.9%) | 181 (19.1%) | 157 (17.1%) | 193 (21.8%) |
| Other | 281 (2.9%) | 39 (4.1%) | 33 (3.6%) | 46 (5.2%) |
| Hispanic (%) | 566 (5.8%) | 117 (12.3%) | 83 (9.0%) | 141 (15.9%) |
| Married/Partnered (%) | 7057 (72.4%) | 565 (59.5%) | 593 (64.5%) | 442 (49.8%) |
| Geographic Region (%) | | | | |
| Northeast | 1555 (16.0%) | 164 (17.3%) | 178 (19.4%) | 172 (19.4%) |
| Midwest | 2680 (27.5%) | 188 (19.8%) | 227 (24.7%) | 165 (18.6%) |
| South | 3756 (38.6%) | 422 (44.4%) | 374 (40.7%) | 411 (46.3%) |
| West | 1746 (17.9%) | 175 (18.4%) | 139 (15.1%) | 139 (15.7%) |
| Education (years) | 12.6 ± 3.0 | 11.3 ± 3.5 | 11.5 ± 3.4 | 10.2 ± 3.7 |
| Education ≥12 years (%) | 7605 (78.0%) | 592 (62.3%) | 592 (64.3%) | 447 (50.4%) |
| Lifestyle | | | | |
| Alcohol use (%) | | | | |
| Non-drinker | 6477 (66.5%) | 716 (75.4%) | 682 (74.1%) | 737 (83.1%) |
| Low/Moderate | 2169 (22.3%) | 117 (12.3%) | 132 (14.3%) | 75 (8.5%) |
| Heavy | 1099 (11.3%) | 117 (12.3%) | 106 (11.5%) | 75 (8.5%) |
| Body Mass Index ($kg/m^2$) | 27.0 ± 4.8 | 27.7 ± 5.7 | 27.9 ± 6.0 | 28.2 ± 5.9 |
| <18.5 | 122 (1.3%) | 17 (1.8%) | 12 (1.3%) | 13 (1.5%) |
| 18.5-<25 | 3394 (34.8%) | 289 (30.4%) | 289 (31.4%) | 259 (29.2%) |
| 25-<30 | 3997 (41.0%) | 375 (39.5%) | 355 (38.6%) | 317 (35.7%) |
| 30-<35 | 1642 (16.8%) | 181 (19.1%) | 163 (17.7%) | 192 (21.6%) |
| ≥35 | 590 (6.1%) | 88 (9.3%) | 101 (11.0%) | 106 (12.0%) |

| | | | | |
|---|---|---|---|---|
| Current Smoker (%) | 1456 (14.9%) | 187 (19.7%) | 175 (19.0%) | 195 (22.0%) |
| Physical Activity (%) | 4927 (50.6%) | 378 (39.8%) | 346 (37.6%) | 247 (27.8%) |
| Time-updated variables in Wave 5 | | | | |
| Married/Partnered (%) | 6887 (70.7%) | 540 (56.8%) | 505 (54.9%) | 404 (45.5%) |
| Alcohol use (%) | | | | |
| Non-drinker | 6819 (70.0%) | 741 (78.0%) | 729 (79.2%) | 754 (85.0%) |
| Low/Moderate | 1959 (20.1%) | 112 (11.8%) | 115 (12.5%) | 72 (8.1%) |
| Heavy | 967 (9.9%) | 97 (10.2%) | 76 (8.3%) | 61 (6.9%) |
| Body Mass Index ($kg/m^2$) | 27.1 ± 5.0 | 27.8 ± 5.6 | 27.8 ± 6.4 | 28.3 ± 6.1 |
| Current Smoker (%) | 1316 (13.5%) | 188 (19.8%) | 155 (16.8%) | 188 (21.2%) |
| Physical Activity (%) | 4818 (49.4%) | 371 (39.1%) | 298 (32.4%) | 196 (22.1%) |

Footnote : Data presented as N (%) or mean±standard deviation where applicable

**Appendix Table 3.2. Sensitivity Analyses for Relationship of Elevated Depressive Symptoms (CES-D≥4) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior Across Multiple G-computation Model Specifications**

| | Outcome | Joint RD [95%SI] | | |
|---|---|---|---|---|
| G-computation Outcome Specification | Model | Overall Cancer | Breast Cancer | Prostate Cancer |
| i) interaction between exposures only | Main Analyses | 0.004 [-0.042,0.052] | -0.018 [-0.042,0.007] | -0.003 [-0.075,0.086] |
| | Bias Analysis with Bias Offset | | | |
| | Low | 0.005 [-0.046,0.054] | -0.018 [-0.040,0.007] | -0.002 [-0.073,0.088] |
| | Moderate | 0.004 [-0.041,0.049] | -0.019 [-0.039,0.007] | -0.006 [-0.078,0.084] |
| | Strong | 0.001 [-0.042,0.055] | -0.019 [-0.039,0.007] | -0.004 [-0.075,0.083] |
| ii) interaction between exposures and exposures-baseline covariates | Main Analyses | -0.001 [-0.045,0.048] | -0.023 [-0.038,-0.006] | -0.008 [-0.077,0.088] |
| | Bias Analysis with Bias Offset | | | |
| | Low | -0.002 [-0.051,0.045] | -0.022 [-0.037,-0.004] | -0.013 [-0.078,0.076] |
| | Moderate | -0.002 [-0.047,0.042] | -0.022 [-0.036,-0.005] | -0.015 [-0.080,0.068] |
| | Strong | -0.005 [-0.050,0.047] | -0.021 [-0.037,-0.005] | -0.013 [-0.078,0.070] |
| iii) interaction between exposures, exposures-baseline covariates, and exposures-current time-varying covariates | Main Analyses | -0.005 [-0.050,0.045] | -0.022 [-0.043,0.002] | -0.017 [-0.079,0.076] |
| | Bias Analysis with Bias Offset | | | |
| | Low | -0.007 [-0.053,0.040] | -0.020 [-0.040,0.005] | -0.021 [-0.082,0.065] |
| | Moderate | -0.007 [-0.053,0.034] | -0.019 [-0.039,0.005] | -0.024 [-0.087,0.059] |
| | Strong | -0.008 [-0.055,0.042] | -0.019 [-0.039,0.007] | -0.022 [-0.080,0.060] |

Abbreviations: SI: simulation interval; RD: risk difference; CES-D: Center for Epidemiological Studies Depression scale score

Footnote: Lower depressive symptoms was the reference in all analyses. Site-specific analyses utilized normal link in final regressions. Bias factors were set to 0.005 and 0.0025 for waves 4 and 5, respectively in the low bias offset model. The moderate and strong bias offset models multiplied each bias factor by 1.5 and 2, respectively.

**Appendix Table 3.3. Sensitivity Analyses Investigating the Relationship of Elevated Depressive Symptoms (CES-D≥3) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior**

| | Joint RD [95%SI] | | |
|---|---|---|---|
| Outcome | Overall Cancer | Breast Cancer | Prostate Cancer |
| N Event/N | 2277/12,502 | 348/7442 | 488/5047 |
| Main Analyses | 0.006 [-0.025,0.036] | -0.024 [-0.039,-0.006] | -0.002 [-0.049,0.052] |
| Bias Analysis with Bias Offset | | | |
| Low | 0.003 [-0.028,0.036] | -0.022 [-0.037,-0.004] | -0.004 [-0.047,0.049] |
| Moderate | 0.003 [-0.029,0.035] | -0.023 [-0.038,-0.003] | -0.004 [-0.049,0.047] |
| Strong Bias | 0.003 [-0.029,0.034] | -0.022 [-0.037,-0.005] | -0.003 [-0.048,0.047] |

Abbreviations: SI: simulation interval; RD: risk difference; CES-D: Center for Epidemiological Studies Depression scale score

Footnote: Lower depressive symptoms was the reference in all analyses. Fitted logistic regression for g-computation included covariates described in text and additionally included interaction terms between exposures, baseline covariates and time-varying covariates. Site-specific analyses utilized normal link in final regressions. Bias factors were set to 0.005 and 0.0025 for waves 4 and 5, respectively in the low bias offset model. The moderate and strong bias offset models multiplied each bias factor by 1.5 and 2, respectively.

**Appendix Table 3.4. Sensitivity Analyses for Relationship of Elevated Depressive Symptoms (CES-D≥3) and Incident Cancer Outcomes Using the Bias Offset for Time-Varying Uncontrolled Confounding due to Sedentary Behavior Across Multiple G-computation Model Specifications**

| | Outcome | Joint RD [95%SI] | | |
|---|---|---|---|---|
| G-computation Outcome Specification | Model | Overall Cancer | Breast Cancer | Prostate Cancer |
| i) interaction between exposures only | Main Analyses | 0.010 [-0.023,0.044] | -0.020 [-0.035,-0.002] | 0.002 [-0.046,0.056] |
| | Bias Analysis with Bias Offset | | | |
| | Low | 0.010 [-0.026,0.044] | -0.020 [-0.036,-0.002] | 0.002 [-0.045,0.059] |
| | Moderate | 0.010 [-0.023,0.045] | -0.019 [-0.036,-0.002] | 0.001 [-0.048,0.059] |
| | Strong | 0.009 [-0.025,0.043] | -0.021 [-0.035,-0.002] | 0.002 [-0.044,0.058] |
| ii) interaction between exposures and exposures-baseline covariates | Main Analyses | -0.001 [-0.045,0.048] | -0.023 [-0.038,-0.006] | -0.008 [-0.077,0.088] |
| | Bias Analysis with Bias Offset | | | |
| | Low | 0.008 [-0.022,0.039] | -0.023 [-0.038,-0.006] | 0.002 [-0.046,0.056] |
| | Moderate | 0.007 [-0.026,0.036] | -0.022 [-0.037,-0.004] | 0.001 [-0.047,0.057] |
| | Strong | 0.006 [-0.024,0.038] | -0.022 [-0.036,-0.005] | -0.001 [-0.048,0.056] |
| iii) interaction between exposures, exposures-baseline covariates, and exposures-current time-varying covariates | Main Analyses | 0.003 [-0.026,0.032] | -0.024 [-0.039,-0.007] | -0.006 [-0.050,0.050] |
| | Bias Analysis with Bias Offset | | | |
| | Low | 0.000 [-0.030,0.032] | -0.022 [-0.036,-0.006] | -0.007 [-0.051,0.045] |
| | Moderate | 0.001 [-0.030,0.034] | -0.023 [-0.039,-0.005] | -0.008 [-0.052,0.045] |
| | Strong | 0.000 [-0.030,0.030] | -0.022 [-0.037,-0.004] | -0.006 [-0.051,0.046] |

Abbreviations: SI: simulation interval; RD: risk difference; CES-D: Center for Epidemiological Studies Depression scale score

Footnote: Lower depressive symptoms was the reference in all analyses. Site-specific analyses utilized normal link in final regressions. Bias factors were set to 0.005 and 0.0025 for waves 4 and 5, respectively in the low bias offset model. The moderate and strong bias offset models multiplied each bias factor by 1.5 and 2, respectively.

# 4. Study 3: Inverse Probability Weighted G-Computation Bias Analysis for Time-Varying Uncontrolled Confounding: Applied to an Investigation of the Link Between Elevated Depressive Symptoms and Cardiovascular Disease

## 4.1 Abstract

**Background:** Although time-varying uncontrolled confounding biases causal inference, methods needed to address its impact in time-varying treatment and confounder settings remain scarce. Here, we demonstrate a new bias modeling method that combines inverse probability of uncontrolled confounder weight (IPUW) with g-computation. The bias analysis method was demonstrated in simulations and used to study the link between depressive symptoms and incident cardiovascular disease (CVD) adjusting for unmeasured time-varying sedentary behavior using data.

**Methods:** The study employed formalization, simulations, and empirical application to demonstrate the utility of combining IPUW with g-computation for bias analysis. The method requires a time-indexed IPUW used to weight a g-computation model conditional on measured confounders. In simulations, we generated large cohorts with repeated measures over two time points to understand the impact of time-varying uncontrolled confounding. We included and excluded the uncontrolled confounders from outcome regression models to obtain true and biased empirical parameters, respectively. These parameters were used in the calculation of potential outcomes for the final marginal structural models to estimate the joint risk difference or odds ratio. In a cohort of 10,212 older adults in 1998 and 2000, we investigated the effect of time-varying persistent elevated depressive symptoms on incident cardiovascular disease (CVD) over ten years, using g-computation models that accounted for demographics and time-varying lifestyle factors. We used IPUW to adjust for simulated time-varying uncontrolled sedentary behavior.

**Results:** In simulations, mis-specified models excluding the time-varying uncontrolled confounders resulted in biased estimates. The combined IPUW and g-computation yielded

similar effect estimates to that of our true joint values. In the cohort of older adults, compared with lower depressive symptoms, the relative risk [95%SI, simulation interval] of joint effect of elevated depressive symptoms on incident CVD was 1.20 [0.91, 1.49]. Similar relationships were observed for combined incident CVD/CVD death outcomes, inclusion of comorbidities in modeling, and conservative definitions of elevated symptoms. Finally, we performed quantitative bias analysis that suggested that only strong levels of time-varying sedentary behavior would change the conclusion.

**Conclusions:** Time-varying uncontrolled confounding can be successfully addressed by the novel use of g-computation combined with IPUW to adjust for time-varying uncontrolled confounding.

## 4.2 Introduction

Bias analysis is critical in addressing the impact of systematic biases including confounding, selection bias, and measurement error[43,52]. Methods of bias analysis includes bias formulas to offset the bias[53]. In addition, bias analysis has been framed as a special case of a missing data problem[52]. Ideally, researchers would want complete data on the exposure, outcome, and all possible confounders both known and theoretically unknown. In reality, data on any critical variable may be mismeasured (i.e., measurement error), partially available for certain individuals (i.e., conditional on a variable for selection bias), or completely missing (i.e., potential uncontrolled confounding). In the latter case, the uncontrolled confounder and its distribution are unknown. Addressing this missing data leads to the complete table, or the distribution of variables if we had the complete data on exposure (X), outcome (Y), covariates (Z), and uncontrolled confounder(s) (U). Understanding how U relates to X, Y, and Z could resolve this problem through the re-weighting of the observed data to obtain the joint distribution including U. Several studies have applied probability of confounder weighting in bias analysis[59,97,98].

Sensitivity analyses using bias analysis has been commonly featured in single time point or baseline studies[49]. With the burgeoning popularity of big data with repeated measurements and interest in causal inference methods, bias analysis is becoming increasingly important in understanding the strong assumptions that come with analyses[46,47,52]. These observational studies as well as meta-analyses are critical in informing public health guidelines and directing resources, especially in scenarios where randomized clinical trials are less feasible[52]. Studies of causal inference including the use of g-computation (g-formula) involves several assumptions including correct model specification and the absence of biases[45]. To address model

specification, researchers have proposed machine learning methods as well as doubly robust estimation (DRE) techniques with weights[99,100]. With DRE methods, inverse probability of treatment weights are used to weight the outcome regression (or Q-model) for g-computation to obtain a correct model specification. Misspecification of the outcome regression but with correct inverse probability of treatment weights would yield a valid result as if either model was originally correctly specified[99]. Weights have traditionally been used in treatment modeling, but they have also been used for informative censoring and mediation analyses[58,101,102].

The use of sensitivity analyses for marginal structural models to assess these biases remain understudied and underutilized[56,58]. To extend the framework of bias analysis, we describe the use of an inverse probability of uncontrolled confounder weight (IPUW) with g-computation. In simulation, we demonstrate the doubly robust estimation properties of IPUW and g-computation modeling to evaluate joint effect estimates. Moreover, we utilize data from a large longitudinal survey of older US adults in a demonstrated application of this novel method. Given noted limitations of confounding control in prior studies, we investigate the relationship of elevated depressive symptoms over two survey waves, with incident cardiovascular disease (CVD) and bias analysis for time-varying uncontrolled sedentary behavior[28].

## 4.3 Methods

### Notation and definitions

Table 4.1 shows the data-generating processes for simulations 3.1-3.3 (Figure 4.1). Simulations included samples of N = 50,000 individuals and were repeated over K = 1000 replications. Data were drawn under a Bernoulli for binary variables. Notation for our directed acyclic graphs (DAGs) and simulations are as follows: let X represent the exposure, Y represent the outcome, Z represent the minimally sufficient set of covariates for measured confounding,

and U represent the uncontrolled confounder. We examined two time points of measurements and each variable is marked with a subscript t for each time point (1 or 2). Figures 4.1A and 4.1B represent the following DAGs used in simulations: two time-varying uncontrolled confounders under a null simulated joint effect, and two time-varying uncontrolled confounders under a null joint effect, where $U_2$ is an intermediate on the path of $X_1$. This represents simulations 3.1 and 3.2, which were simulated on a linear scale to investigate a joint and null risk difference. Finally, simulation 3.3 investigated two time-varying uncontrolled confounders under a non-null simulated joint effect and simulated on the logit scale (Figure 4.1A).

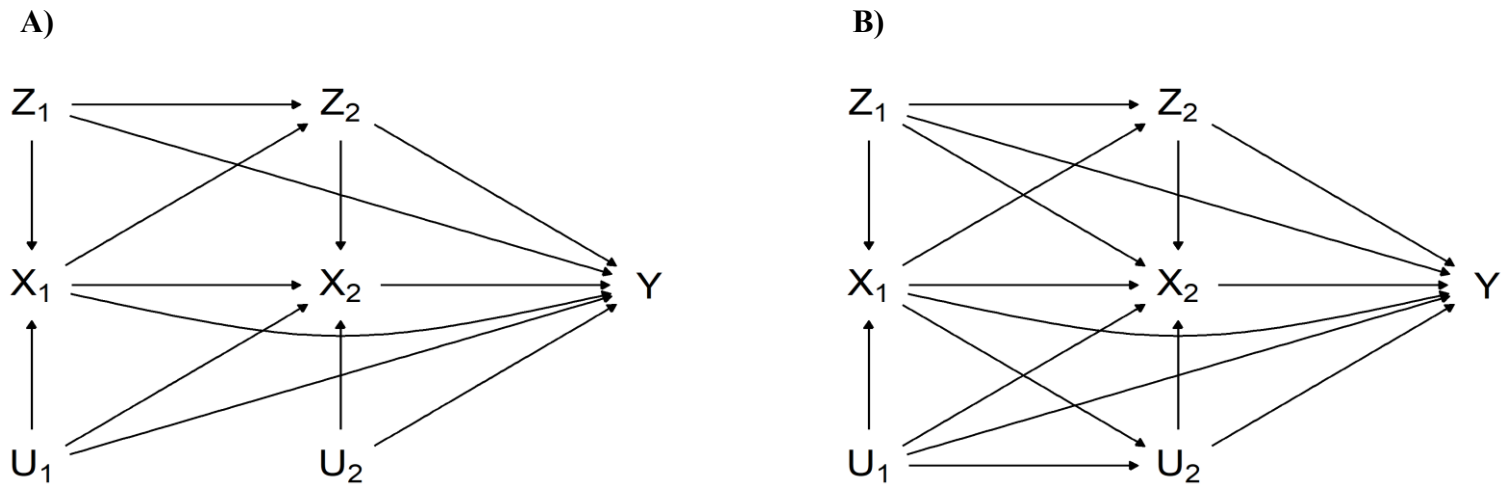
**Table 4.1. Parameters for Simulations 3.1-3.3**

| Variable | Description | Simulation 3.1:<br><br>Time-varying exposures with 2 uncontrolled confounders on the linear scale | Simulation 3.2:<br><br>Time-varying exposures with 2 uncontrolled confounders and a mediator on the linear scale | Simulation 3.3:<br><br>Time-varying exposures with 2 uncontrolled confounders on the logit scale |
|---|---|---|---|---|
| $U_1$ | **Uncontrolled confounder at time = 1** | 0.4 | 0.4 | 0.6 |
| $Z_1$ | **Covariate at time = 1** | 0.5 | 0.5 | 0.4 |
| $X_1$ | **Exposure at time = 1** | $0.4 + 0.1Z_1 + R_{U1}U_1$ | $0.4 + 0.1Z_1 + R_{U1}U_1$ | $\text{expit}(\log(0.50) + \log(1.50)Z_1 + R_{U1}U_1)$ |
| $U_2$ | **Uncontrolled confounder at time = 2** | 0.4 | $0.20 + 0.2U_1 + 0\times X_1$ | 0.6 |
| $Z_2$ | **Covariate at time = 2** | $0.2 + 0.1Z_1 + 0\times X_1$ | $0.20 + 0.10Z_1 + 0\times X_1$ | $\text{expit}(\log(0.50) + \log(1.50)Z_1 + \log(2.00)X_1)$ |
| $X_2$ | **Exposure at time = 2** | $0.15 + 0.2X_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $0.15 + 0.2X_1 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $\text{expit}(\log(0.40) + \log(2.50)X_1 + \log(1.50)Z_2 + R_{U1}U_1 + R_{U2}U_2)$ |
| $Y$ | **Outcome after time = 2** | $0.1 + 0\times X_1 + 0\times X_2 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $0.1 + 0\times X_1 + 0\times X_2 + 0.1Z_1 + 0.1Z_2 + R_{U1}U_1 + R_{U2}U_2$ | $\text{expit}(\log(0.05) + \log(1.10)X_1 + \log(1.20)X_2 + \log(1.20)X_1X_2 + \log(1.20)Z_1 + \log(1.20)Z_2 + R_{U1}U_1 + R_{U2}U_2)$ |

Abbreviations: expit: inverse logit; $R_{Ut}$: the parameter linking $U_t$ and the specified variables; $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome; $Z_t$: covariates at time t

Footnote: To demonstrate a null effect with simulations 3.1 and 3.2 across all bias analysis models, each indirect relationship (i.e., $X_1 \rightarrow Z_2$ and $X_1 \rightarrow U_2$) was also set to null. All variables were generated as binary variables.

**Figure 4.1. Directed Acyclic Graphs of Time-Varying Exposures Depicting A) Only Two Time-Varying Uncontrolled Confounders and B) Two Time-Varying Uncontrolled Confounders and an Unmeasured Mediator**




Abbreviations: $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 2; $Z_t$: covariates at time t

**G-computation**

Detailed tutorials for g-computation have been previously described elsewhere[45,65]. In our simulations, we assumed that we met all assumptions except for conditional exchangeability as this assumption was under investigation here.

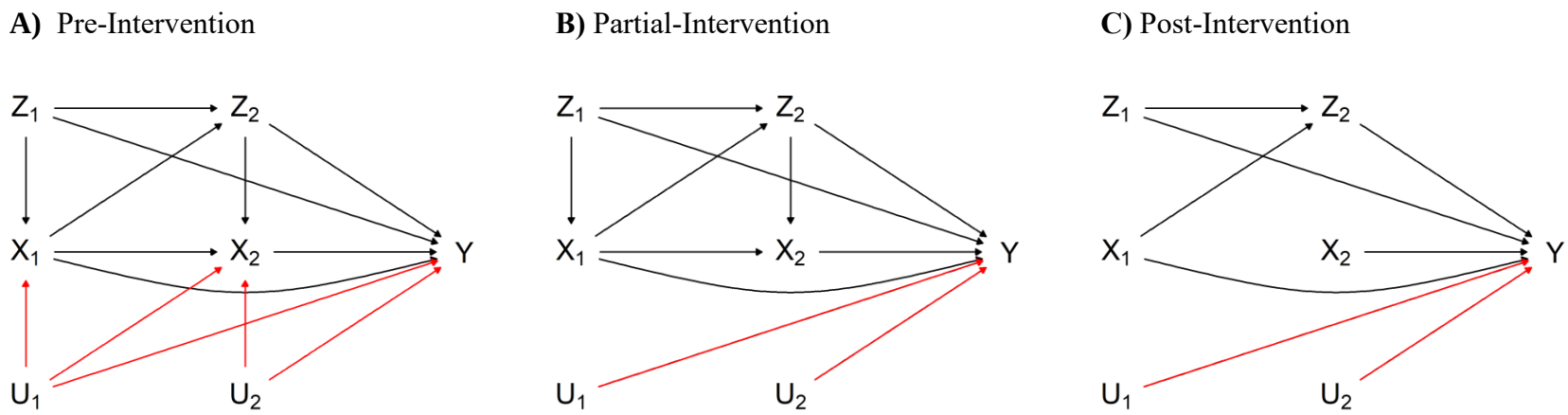
We first modeled the descendants of X (i.e., time-varying covariates) and the outcome using the appropriate variables as dictated by our DAG. As data were simulated, our models were correctly specified including the addition of interaction terms where needed and the appropriate regression link. In our simulations, potential outcomes were calculated at each time point to represent hypothetical treatment or no treatment using the regression parameters obtained in the previous step. The joint risk difference (RD) was obtained by contrasting the potential outcome for hypothetical treatment at both time points and no treatment ever. The joint odds ratio (OR) was similarly obtained by contrasting the potential outcomes using the corresponding odds ratio equation. We took the distribution of the mean value per replication for all 1000 replications. The median and $2.5^{th}$ and $97.5^{th}$ percentiles of estimates from all replications were considered as our final effect estimate and 95% simulation interval (SI). The bias and root mean square error were computed for each effect estimate.

In the outcome regressions, we first included all necessary variables including $U_t$ to represent our 'True' model and estimate. In this scenario, $U_t$ is in fact known and measured. Next, we excluded $U_t$ in the relevant regression models and calculations to represent the 'Biased' model and estimate. These latter models were purposely mis-specified, in violation of the conditional exchangeability assumption, and has uncontrolled confounding due to the exclusion of $U_t$ from Q-models. Simulations were repeated to vary the parameters relating $U_t$ to $X_t$ and Y, respectively to individually understand the impact of uncontrolled confounding on our results.

**Bias Analysis with Weighting**

We examined the doubly robust estimation properties in calculating IPUW combined with g-computation for our bias analysis. Figure 4.2A (or Figure 4.1A) represents the pre-intervention DAG depicting the theorized data generation process. The application of the IPUW on the DAG is shown in Figure 4.2B, where the arrows from $U_t$ are removed, while other observed paths are preserved. This is achieved by weighting the Q-model with the IPUW and using those weighted regression parameters in subsequent g-computation steps. Figure 4.2C reflects the post-intervention DAG where arrows going into $X_t$ are removed so that the appropriate paths can be correctly identified without bias. If $U_t$ were known, Figure 4.2C would represent a typical post-intervention DAG with g-computation.

**Figure 4.2. Directed Acyclic Graphs Depicting A) Pre-Intervention, B) Partial-Intervention After Inverse Probability of Uncontrolled Confounder Weight, and C) the Post-Intervention after G-computation**




Abbreviations: $U_t$: uncontrolled confounder at time t; $X_t$: exposure at time t; Y: outcome after time = 2; $Z_t$: covariates at time t

The IPUW resembles stabilized inverse probability of treatment weights. As inverse probability of treatment weighting would be $\frac{P(X_1,X_2)}{P(X_1,X_2|\ U_1,U_2\ )}$ for Figure 4.1A for example, rearranging these probabilities would yield an IPUW of $\frac{P(U_1,U_2)}{P(U_1,U_2|\ X_1,X_2\ )}$. $Z_t$ could also be included in the IPUW denominator depending on whether it is independent of $U_t$. We used logistic regression to model the probability of $U_t$ ($U_t = 1$) for each time point as our numerator. We examined two sets of denominators: with and without $Z_t$ covariates. Using logistic regression, we modeled the probability of $U_t$ given the exposures, and/or $Z_t$ as the denominator. The division of the numerator and denominator reflected the IPUW and was assigned based on the actual value of $U_t$. IPUW at each time point were multiplied together to represent the final stabilized IPUW used in analyses. The mean of the stabilized IPUW were 1.00 as expected.

We used the following steps to implement the bias weighting in simulations:

Step 1: Calculate the numerator and denominator of the IPUW accordingly for each time point using logistic regression. The final stabilized IPUWs were $\frac{P(U_1)}{P(U_1|X_1,X_2,Z_1,Z_2)} \times \frac{P(U_2)}{P(U_2|U_1,X_1,X_2,Z_1,Z_2)}$ or $\frac{P(U_1)}{P(U_1|X_1,X_2)} \times \frac{P(U_2)}{P(U_2|U_1,X_1,X_2)}$ if $Z_t$ were excluded from the IPUW.

Step 2: Fit the biased outcome regression weighted by the IPUW using the appropriate regression link and adjusting for the observed covariates. Note that the regression was intentionally mis-specified because $U_1$ and $U_2$ were excluded from the regression.

Step 3: Use the parameters from the original g-computation and marginal distributions to re-simulate a plasmode dataset for the exposures and all covariates. Use the weighted regression parameters (from Step 2) to re-simulate Y (i.e., the outcome weighted by the time-varying uncontrolled confounding).

Step 4: Run the final marginal structural model for the desired effect estimate with the simulated data for each replication. The median, and 2.5$^{th}$ and 97$^{th}$ percentiles of all replications then reflect the effect estimate and 95% simulated interval.

### 4.3.1 Illustration with the Health and Retirement Study

We used data from the Health and Retirement Study (HRS) to investigate the joint relationship of elevated depressive symptoms with incident cardiovascular disease. The source HRS data has been previously described[84]. For this illustration, we utilized the RAND summarized data files (2018 v1) to identify the cohort of age-eligible and CVD-free participants in 1998 and 2000. The data from 1998 and 2000 represents survey waves 4 and 5, respectively and were the earliest periods of data harmonization in the HRS. For the incident CVD outcome, participants were followed for up to ten years (2010) or were censored for death or lost to follow-up. We excluded participants for missing data on depressive symptoms or answering less than half of the survey, and missing data on any covariates over the study period. The final cohort comprised 10,212 older adults (Appendix Figure 4.1).

The primary exposure were the responses from the Center for Epidemiological Studies Depression scale (CES-D) short form survey which consisted of eight scaled questions. Elevated depressive symptoms were defined as a CES-D score $\geq$4[86-88]. Incident CVD was defined as a self-reported positive response to having heart disease (myocardial infarction, coronary heart disease, angina, congestive heart failure or other heart problems) or stroke (or a possible stroke or transient ischemic attack), as defined by the RAND files. Derived RAND Fat files were used to quantify causes of reported heart disease. In consideration of prior studies, we used the following baseline and time-varying covariates in modeling: baseline age, gender, race (white/non-white), ethnicity, high-school equivalent education and time-varying

married/partnered status, current alcohol status (drinker/non-drinker, defined by the number of drinks per day when drinking), current smoking status (smoker/non-smoker), body mass index and performing vigorous physical activity at least three times/week[33,36,90,103] (Appendix Figure 4.2A).

**Analyses with the Health and Retirement Study**

Data were summarized as mean ± standard deviation or frequency (percentage, %), as applicable. Clinical characteristics were stratified by depressive symptoms (CES-D <4 vs ≥ 4) in at baseline wave 4, as well as by level of depressive symptoms across both wave 4 and 5.

We performed g-computation in the fitting of marginal structural models as described above to obtain the joint relative risk (RR) and joint RD. In summary, we fit regression models reflecting the data generating process with a logit link for each time-varying covariate with an interaction term for the prior exposure and covariate of interest. The outcome regression model was also fit with a logit link and two-way interaction terms across all exposures and covariates. Data were simulated for 1000 replications. The final marginal structural model represented the joint estimate of persistent elevated depressive symptoms (reference: no elevated depressive symptoms in both time points) with incident CVD outcomes.

We performed several sensitivity analyses before quantitative bias analysis to address other potential biases in the analyses. We explored potential exposure misclassification in defining elevated depressive symptoms as a CES-D score ≥3[88]. We evaluated additional outcome models for g-computation that included various combinations of two-way interaction terms to evaluate model misspecification. We included time-varying comorbidities of diabetes and high blood pressure given the relationship between metabolic syndrome and cardiovascular disease[104]. We also investigated short-term and long-term relationships with CVD after two years or the

maximum of follow-up (end of study period, 2018), respectively. Finally, we examined a combined outcome of incident CVD and reported death due to heart, circulatory and blood conditions as CVD death.

As to investigate potential uncontrolled confounding due to time-varying sedentary behavior, we applied weighting in bias analysis as described above (Appendix Figure 4.2B). Sedentary behavior was unavailable from the HRS during this study period yet has been identified potential risk factor for each depressive symptoms and CVD[41,105,106]. Furthermore, we assumed that depressive symptoms would additionally inform future sedentary behavior as a mediating effect[107]. The inputs and bias parameters to repair this joint distribution were informed by prior studies, available data from the HRS on lack of physical activity and theorized the prevalence of sedentary behavior and relationships to other variables. First, we calculated the probability of sedentary behavior using an expit function and informed bias parameters for the exposures, outcome and covariates ($P(U_4|\text{CES-D}_4,\text{CES-D}_5,\text{CVD},V_4,Z_4,Z_5)$ and $P(U_5|U_4,\text{CES-D}_4,\text{CES-D}_5,\text{CVD},V_4,Z_4,Z_5)$, for each time point). The relationships between $U_t$ to exposures and outcome were varied in multiple bias analysis models. These probabilities were used to simulate a binary variable representing sedentary behavior at each time point. The IPUW numerator and denominator were calculated with logistic regression to output the predicted probabilities for each time point. The numerator was the marginal probability of sedentary behavior or modified as $P(U_5|\text{CES-D}_4)$ to preserve the mediating paths due to $U_5$. The denominator was the probability of sedentary behavior as a function of the exposures and covariates. The IPUW was calculated for each observation based on actual value of the sedentary behavior, per time point. Each IPUW per time point were multiplied together for the final IPUW used in analyses. The outcome regression model was fit with all exposures, covariates, and two-way interactions as in the main

analyses and was weighted by the stabilized IPUW. Using the marginal distribution and the parameters obtained from the intermediate regression models from the main analyses, the exposures and covariates were re-simulated. The bias-weighted outcome was simulated using the weighted empirical parameters. The simulation was repeated for K = 1000 replications. The final marginal structural model included the simulated exposures and outcome with the appropriate link. Estimates were summarized per replication to obtain the median effect estimate and the 95% SI.

We chose biasing scenarios that varied the relationships between exposures, outcome, and uncontrolled confounders to perform our quantitative bias analysis for time-varying uncontrolled confounding. We first assumed that $U_5$ acted as a mediator and confounder. Across a range of bias parameters (log(0.60) to log(2.60) for the log(OR)), we assumed that the relationships from $U_t$ to each exposure and outcome were of the same magnitude at both time points. In these analyses, we included and excluded $Z_t$ from the denominator of the IPUW to demonstrate the model specification properties. Moreover, we further incrementally changed the relationships of $U_5$ while holding the relationship at wave 4 at a constant value (log(1.40)), given the unique role of $U_5$ as a mediator and confounder. Finally, we repeated the series of bias analysis methods under the assumption that time-varying sedentary behavior was only a confounder (similar to Figure 4.1A). Therefore, the numerator of the IPUW was modeled as the marginal probabilities of the uncontrolled confounder at each time point.

Analyses were performed using SAS 9.4 (Cary, NC), and R (R version 4.0.0 and R Studio version 1.2.5042). Data were not missing in simulations and complete case was used in HRS analyses. This study was exempt from institutional review board review given that it is public use and de-identified data.

## 4.4 Results

### 4.4.1 Simulations Demonstrating the Bias Weight

#### Simulation 3.1 & 3.2

In simulation 3.1, the true joint RD [95%SI] was null at 0.000 [-0.009, 0.009] (Table 4.2). The observed biased joint RD differed from the null as the strength of $U_t$ increased, including the result of 0.027 [0.018, 0.037] when both uncontrolled confounders had relationships of 0.10 to the exposures and outcome. When modeling an IPUW for each $U_t$, the subsequent joint RD was again null, with albeit slightly wider confidence intervals. The inclusion or exclusion of $Z_t$ within the IPUW specification both yielded similar results as the g-computation was correctly specified and already included $Z_t$.

Again, simulation 3.2 had a true joint RD [95%SI] of approximately 0.000 [-0.009, 0.009], or null (Table 4.3). In this simulation, there was added complexity in that the uncontrolled confounder, $U_2$ acted as a mediator and had a parallel role to that of $Z_t$. As $U_t$ incrementally increased in value, the amount of bias, or how far the joint RD deviated from the null also increased. Likewise, having a strong confounder at an earlier time point, resulted in biased results. When using the IPUW, we observed similar conclusions to that of simulation 3.1. The use of an IPUW when modeling the outcome for g-computation resulted in a null joint RD and this was observed with or without $Z_t$ in the IPUW specification.

**Table 4.2. Simulation Results for the Joint Risk Difference in Simulation 3.1 that Included $Z_t$ and Excluded $Z_t$ in the IPUW**

| $R_{U1}$ | $R_{U2}$ | Model | Model | Joint RD [95%SI] | Bias | RMSE |
|---|---|---|---|---|---|---|
| 0.025 | 0.025 | 1 | True | 0.000 [-0.009,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.002 [-0.007,0.011] | 0.002 | 0.005 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.013,0.013] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.000 [-0.013,0.014] | 0.000 | 0.007 |
| 0.05 | 0.05 | 2 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.007 [-0.003,0.017] | 0.007 | 0.008 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.013,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.000 [-0.013,0.015] | 0.000 | 0.007 |
| 0.10 | 0.10 | 3 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.027 [0.018,0.037] | 0.027 | 0.028 |
| | | | Biased+IPUW with $Z_t$ | 0.001 [-0.014,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.015] | 0.001 | 0.007 |
| 0.025 | 0.10 | 4 | True | 0.000 [-0.009,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.011 [0.001,0.020] | 0.011 | 0.012 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.014,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.000 [-0.013,0.015] | 0.000 | 0.007 |
| 0.10 | 0.025 | 5 | True | 0.000 [-0.010,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.019 [0.009,0.028] | 0.019 | 0.020 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.013,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.015] | 0.000 | 0.008 |

Abbreviations: IPUW: inverse probability of uncontrolled confounder weight; RD: risk difference; RMSE: root mean square error; $R_{Ut}$: risk difference for the relationship between $U_t$ and the respective variables; SI: simulation interval; $Z_t$; covariates at time 1 and 2

Footnote: The true joint effect is set to 0.00 in simulations. True models represent results from the correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where $U_t$ were excluded. Biased + IPUW models represent the the incorrectly specified g-computation regression, weighted by the IPUW.

**Table 4.3. Simulation Results for the Joint Risk Difference in Simulation 3.2 that Included $Z_t$ and Excluded $Z_t$ in the IPUW**

| $R_{U1}$ | $R_{U2}$ | Model | Model | Joint RD [95%SI] | Bias | RMSE |
|---|---|---|---|---|---|---|
| 0.025 | 0.025 | 1 | True | 0.000 [-0.009,0.009] | 0.000 | 0.004 |
| | | | Biased | 0.002 [-0.007,0.011] | 0.002 | 0.005 |
| | | | Biased+IPUW with $Z_t$ | 0.001 [-0.013,0.014] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.014] | 0.000 | 0.007 |
| 0.05 | 0.05 | 2 | True | 0.000 [-0.009,0.009] | 0.000 | 0.004 |
| | | | Biased | 0.008 [-0.001,0.017] | 0.008 | 0.009 |
| | | | Biased+IPUW with $Z_t$ | 0.001 [-0.014,0.014] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.014] | 0.001 | 0.007 |
| 0.10 | 0.10 | 3 | True | 0.000 [-0.009,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.031 [0.022,0.040] | 0.031 | 0.031 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.014,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.016] | 0.001 | 0.008 |
| 0.025 | 0.10 | 4 | True | 0.000 [-0.009,0.009] | 0.000 | 0.005 |
| | | | Biased | 0.010 [0.002,0.020] | 0.010 | 0.011 |
| | | | Biased+IPUW with $Z_t$ | 0.001 [-0.013,0.015] | 0.001 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.015] | 0.001 | 0.007 |
| 0.10 | 0.025 | 5 | True | 0.000 [-0.009,0.009] | 0.000 | 0.004 |
| | | | Biased | 0.021 [0.012,0.029] | 0.020 | 0.021 |
| | | | Biased+IPUW with $Z_t$ | 0.000 [-0.014,0.015] | 0.000 | 0.007 |
| | | | Biased+IPUW without $Z_t$ | 0.001 [-0.013,0.015] | 0.001 | 0.007 |

Abbreviations: IPUW: inverse probability of uncontrolled confounder weight; RD: risk difference; RMSE: root mean square error; $R_{Ut}$: risk difference for the relationship between $U_t$ and the respective variables; SI: simulation interval; $Z_t$; covariates at time 1 and 2

Footnote: The true joint effect is set to 0.00 in simulations. True models represent results from the correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where $U_t$ were excluded. Biased + IPUW models represent the the incorrectly specified g-computation regression, weighted by the IPUW.

**Simulation 3.3**

Finally, simulation 3.3 had a true joint OR of approximately 1.63 (Table 4.4). Under weaker levels of confounding, the bias estimate was very similar to the true joint OR. Across incrementally stronger effects of time-varying uncontrolled confounding had more biased results. Having both confounders with extreme strengths of 3.00 yielded a joint OR [95%SI] of 2.98 [2.80, 3.17]. The use of an IPUW in combination with g-computation gave similar results to that of the true joint OR. Notably, situations of very extreme confounding, such as the strength of 3.00, the combined IPUW gave a slightly lower magnitude of estimate than expected. Finally, we observed similar conclusions when including or excluding $Z_t$ from the IPUW specification.

**Table 4.4. Simulation Results for the Joint Odds Ratio in Simulation 3.3 that Included $Z_t$ and Excluded $Z_t$ in the IPUW**

| $R_{U1}$ | $R_{U2}$ | Model | Model | Joint OR [95%SI] | Bias | RMSE |
|---|---|---|---|---|---|---|
| 1.20 | 1.20 | 1 | True | 1.63 [1.50,1.77] | -0.001 | 0.067 |
| | | | Biased | 1.66 [1.53,1.80] | 0.031 | 0.075 |
| | | | Biased+IPUW with $Z_t$ | 1.62 [1.44,1.84] | -0.003 | 0.103 |
| | | | Biased+IPUW without $Z_t$ | 1.62 [1.44,1.84] | -0.003 | 0.103 |
| 1.50 | 1.50 | 2 | True | 1.63 [1.52,1.75] | 0.000 | 0.061 |
| | | | Biased | 1.78 [1.67,1.92] | 0.155 | 0.168 |
| | | | Biased+IPUW with $Z_t$ | 1.62 [1.46,1.81] | -0.004 | 0.089 |
| | | | Biased+IPUW without $Z_t$ | 1.63 [1.46,1.82] | -0.002 | 0.089 |
| 2.00 | 2.00 | 3 | True | 1.63 [1.53,1.75] | 0.000 | 0.056 |
| | | | Biased | 2.10 [1.98,2.25] | 0.470 | 0.476 |
| | | | Biased+IPUW with $Z_t$ | 1.61 [1.47,1.79] | -0.019 | 0.084 |
| | | | Biased+IPUW without $Z_t$ | 1.61 [1.47,1.79] | -0.014 | 0.083 |
| 2.50 | 2.50 | 4 | True | 1.63 [1.53,1.76] | 0.003 | 0.057 |
| | | | Biased | 2.51 [2.35,2.68] | 0.881 | 0.885 |
| | | | Biased+IPUW with $Z_t$ | 1.59 [1.45,1.76] | -0.037 | 0.089 |
| | | | Biased+IPUW without $Z_t$ | 1.60 [1.45,1.77] | -0.029 | 0.086 |
| 3.00 | 3.00 | 5 | True | 1.64 [1.53,1.75] | 0.002 | 0.056 |
| | | | Biased | 2.98 [2.80,3.17] | 1.348 | 1.352 |
| | | | Biased+IPUW with $Z_t$ | 1.57 [1.43,1.73] | -0.058 | 0.098 |
| | | | Biased+IPUW without $Z_t$ | 1.58 [1.44,1.74] | -0.048 | 0.093 |
| 1.50 | 3.00 | 6 | True | 1.64 [1.53,1.75] | 0.003 | 0.055 |
| | | | Biased | 2.10 [1.97,2.22] | 0.463 | 0.468 |
| | | | Biased+IPUW with $Z_t$ | 1.60 [1.46,1.77] | -0.024 | 0.082 |
| | | | Biased+IPUW without $Z_t$ | 1.61 [1.47,1.79] | -0.020 | 0.081 |
| 3.00 | 1.50 | 7 | True | 1.63 [1.52,1.76] | 0.001 | 0.061 |
| | | | Biased | 2.58 [2.41,2.76] | 0.948 | 0.953 |
| | | | Biased+IPUW with $Z_t$ | 1.60 [1.45,1.79] | -0.025 | 0.092 |
| | | | Biased+IPUW without $Z_t$ | 1.61 [1.46,1.80] | -0.016 | 0.090 |

Abbreviations: IPUW: inverse probability of uncontrolled confounder weight; OR: odds ratio; RMSE: root mean square error; $R_{Ut}$: parameter for the relationship between $U_t$ and the respective variables; SI: simulation interval; $Z_t$; covariates at time 1 and 2

Footnote: The true joint effect is set to 1.63 in simulations. True models represent results from the correctly specified g-computation models. Biased models represent the incorrectly specified g-computation model where $U_t$ were excluded. Biased + IPUW models represent the the incorrectly specified g-computation regression, weighted by the IPUW.

### 4.4.2 Clinical Illustration with the Health and Retirement Study

In this cohort of 10,212 participants, the mean±SD age was 64±9 years old and included 62% females (Table 4.5). There were 1305 (13%) participants who had elevated depressive symptoms with a CES-D score ≥4 at baseline. Over two waves, 81% of participants did not have elevated depressive symptoms, while 589 (5.8%) participants experienced persistent elevated depressive symptoms over both waves (Appendix Table 4.1). At baseline, there were a greater percentage of females, non-whites, less educated persons, more non-drinkers, and current smokers in the elevated depressive symptoms group. Similar observations about baseline characteristics were noted among the persistent elevated depressive symptoms group, or those with higher scores across both waves.

**Table 4.5. Cohort Descriptives of 10,212 Health and Retirement Study Participants**

**Stratified by Baseline Level of Depressive Symptoms**

| Characteristic | Total | No elevated depressive symptoms: CES-D <4 | Elevated depressive symptoms: CES-D ≥4 |
|---|---|---|---|
| N (%) | 10,212 | 8907 (87.2%) | 1305 (12.8%) |
| Age (years) | 63.9 ± 9.0 | 63.9 ± 8.8 | 63.4 ± 9.3 |
| 50-<60 | 3812 (37.3%) | 3019 (36.7%) | 294 (41.1%) |
| 60-<70 | 3702 (36.3%) | 3052 (37.1%) | 241 (33.7%) |
| 70-<80 | 2085 (20.4%) | 1713 (20.8%) | 136 (19.0%) |
| ≥80 | 613 (6.0%) | 453 (5.5%) | 45 (6.3%) |
| Gender (% Female) | 6358 (62.3%) | 4889 (59.4%) | 504 (70.4%) |
| Race (%) | | | |
| White | 8545 (83.7%) | 7023 (85.3%) | 556 (77.7%) |
| African American | 1329 (13.0%) | 965 (11.7%) | 130 (18.2%) |
| Other | 338 (3.3%) | 249 (3.0%) | 30 (4.2%) |
| Hispanic (%) | 780 (7.6%) | 507 (6.2%) | 103 (14.4%) |
| Married/Partnered (%) | 7177 (70.3%) | 6001 (72.9%) | 428 (59.8%) |
| Geographic Region (%) | | | |
| Northeast | 1695 (16.6%) | 1309 (15.9%) | 128 (17.9%) |
| Midwest | 2653 (26.0%) | 2238 (27.2%) | 143 (20.0%) |
| South | 4011 (39.3%) | 3155 (38.3%) | 310 (43.3%) |
| West | 1843 (18.1%) | 1527 (18.5%) | 135 (18.9%) |
| Education (years) | 12.4 ± 3.1 | 12.7 ± 3.0 | 11.5 ± 3.5 |
| Education ≥12 years (%) | 7787 (76.3%) | 6570 (79.8%) | 464 (64.8%) |
| Lifestyle Factors | | | |
| Alcohol use (%) | | | |
| Non-drinker | 6820 (66.8%) | 5335 (64.8%) | 519 (72.5%) |
| Low/Moderate | 2109 (20.7%) | 1867 (22.7%) | 89 (12.4%) |
| Heavy | 1283 (12.6%) | 1035 (12.6%) | 108 (15.1%) |
| Body Mass Index ($kg/m^2$) | 27.1 ± 5.0 | 26.9 ± 4.8 | 27.3 ± 5.6 |
| <18.5 | 133 (1.3%) | 98 (1.2%) | 14 (2.0%) |
| 18.5-<25 | 3552 (34.8%) | 2918 (35.4%) | 242 (33.8%) |
| 25-<30 | 4128 (40.4%) | 3399 (41.3%) | 268 (37.4%) |
| 30-<35 | 1711 (16.8%) | 1332 (16.2%) | 127 (17.7%) |
| ≥35 | 688 (6.7%) | 490 (5.9%) | 65 (9.1%) |
| Current Smoker (%) | 1695 (16.6%) | 1265 (15.4%) | 152 (21.2%) |
| Physical Activity (%) | 5064 (49.6%) | 4305 (52.3%) | 309 (43.2%) |
| Pre-existing conditions (%) | | | |
| High Blood Pressure | 4071 (39.9%) | 3151 (38.3%) | 310 (43.3%) |
| Diabetes | 1036 (10.1%) | 766 (9.3%) | 82 (11.5%) |

Footnote : Data presented as N(%) or mean±standard deviation, where applicable

Over a maximum follow-up of approximately ten years (median: 10.1 years), there were 2694 incident CVD events for an event rate [95%CI] of 30.7 [29.5, 31.8] per 1000 person-years. Subjects with elevated depressive symptoms at baseline and persistently over both time points had higher crude event rates compared with other groups. Most reported CVD events were due to heart disease. Among heart disease reports, the most common cause was angina, followed by myocardial infarction and congestive heart failure. Using g-computation to investigate incident CVD, elevated depressive symptoms (CES-D ≥4) compared with lower depressive symptoms at both time points demonstrated a joint RR [95% SI] and RD [95%SI] of 1.20 [0.91, 1.49] and 0.051 [-0.024, 0.122], respectively (Table 4.6).

We performed a series of additional analyses for our main results. First, we examined combined incident CVD/CVD death and observed 3113 events with an event rate [95% confidence interval (CI)] of 35.4 [34.2, 36.7] per 1000 person-years. The observed joint RR and RD for the incident CVD/CVD death outcome were consistent with the main analysis (Table 4.6). In a sub-cohort of participants with data on time-updated diabetes and elevated blood pressure, we observed similar results to that of our main analyses. Moreover, we tested varying levels of two-way interactions in the outcome specification models. Including only an exposure interaction term suggested slightly greater magnitudes of effects. To evaluate a more immediate relationship, we shorten the follow-up period to two years after the last wave and observed 653 CVD events. In this shorter analysis, persistent elevated depressive symptoms had a joint RR [95%SI] and RD [95%SI] of 1.23 [0.62, 1.98] and 0.014 [-0.024, 0.058], respectively compared with lower symptoms at both time points. Evaluation of the long-term relationship over a maximum of 18 years demonstrated a joint RR [95%SI] and RD [95%SI] of 1.08 [0.87, 1.28]

and 0.029 [-0.47, 0.102], respectively. Finally, we evaluated a definition of elevated depressive symptoms as CES-D≥3, rather than the original cutoff at 4 and observed similar results.

**Table 4.6. Joint Relationship of Elevated Depressive Symptoms (CES-D ≥4) and Incident CVD over Ten Years in 10,212 Health and Retirement Study Participants and Under Sensitivity Analyses**

| | Joint Relative Risk [95%SI] | Joint Risk Difference [95%SI] |
|---|---|---|
| **Analyses** | | |
| Main Analysis, without bias analysis | 1.20 [0.91,1.49] | 0.051 [-0.024,0.122] |
| **Sensitivity Analyses** | | |
| Outcome definition of Incident CVD/CVD Death | 1.14 [0.88,1.39] | 0.042 [-0.035,0.115] |
| Inclusion of diabetes & high blood pressure * | 1.16 [0.90,1.49] | 0.041 [-0.026,0.123] |
| Outcome Specification | | |
| i)interaction between exposures | 1.28 [0.99,1.56] | 0.070 [-0.003,0.141] |
| ii)interaction between exposures and exposures-baseline covariates | 1.22 [0.93,1.50] | 0.056 [-0.019,0.128] |
| iii)interaction between exposures, exposures-baseline covariates, and exposures-current time-varying covariates | 1.20 [0.91,1.47] | 0.051 [-0.022,0.117] |
| Short-term follow-up (two years) | 1.23 [0.62,1.98] | 0.014 [-0.024,0.058] |
| Long-term follow-up (maximum: 18 years) | 1.08 [0.87,1.28] | 0.029 [-0.047,0.102] |
| Exposure definition of CES-D ≥3 | 1.14 [0.95,1.32] | 0.035 [-0.014,0.079] |

Abbreviations: CES-D: Center for Epidemiological Studies Depression scale score; CVD: cardiovascular disease; SI: simulation interval

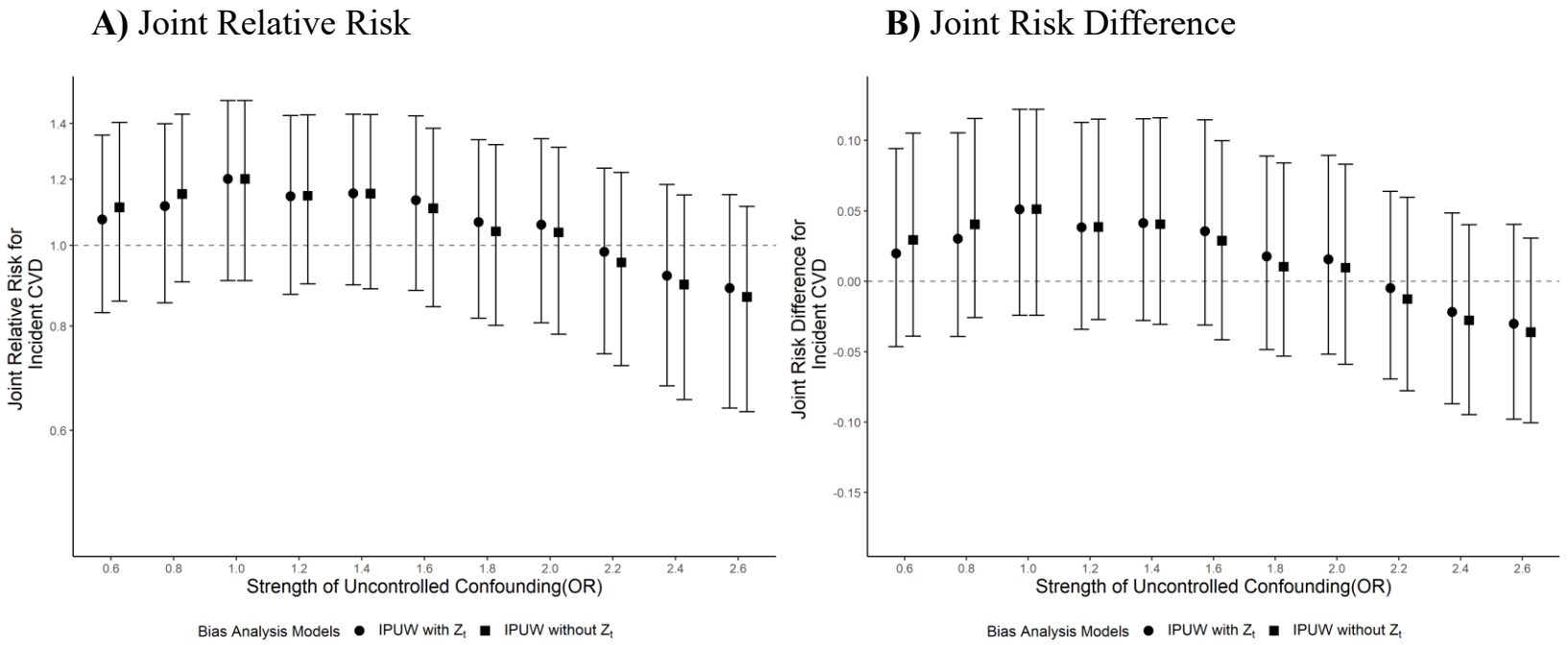
Footnote: Lower depressive symptoms was the reference in all analyses. Main analysis represents the results from the g-computation without bias analysis. Fitted logistic regression for g-computation included covariates described in text and additionally included interaction terms between exposures, baseline covariates and time-varying covariates.

***** N = 10,201 persons with available data on diabetes and high blood pressure

In bias analysis, we performed a series of weighted regression models to understand the impact of time-varying uncontrolled sedentary behavior. Under the assumption that $U_5$ acted as both a mediator and confounder, the relative risk and risk difference gradually declined with increasing bias parameters (Figure 4.3A and B). A protective relationship was merely suggested in the presence of very strong bias parameters for both effect estimates. Under modest parameters, the results from the bias analysis suggested robust and similar results to our main analyses. Moreover, these observations remained the same across all iterations of bias analysis models: inclusion and exclusion of $Z_t$ from the IPUW denominator. Conversely, we observed different results had we assumed that $U_5$ acted as only a confounder (Figure 4.4A and B). While we could make similar conclusions to that of our main results under modest parameters for both effect estimates, results were magnified under much stronger parameters. Specifically, parameters of log(2.2) and greater were sufficient to move both effect estimates and respective CI under the null and suggest a protective relationship of joint elevated depressive symptoms and incident CVD. Repeating all analyses using a longer period of follow-up suggested similar conclusions (Appendix Figure 4.3).

At wave 5, we assumed that sedentary behavior had a unique role in acting as a mediator and a confounder. Under this assumption, incrementally changing the relationship of the uncontrolled sedentary behavior at wave 5 suggested robust results even in the presence of strong bias parameters for both effect estimates (Appendix Figure 4.4A and B). However, under the assumption that sedentary behavior only acted as a confounder in wave 5, bias analysis results were robust in modest strengths of bias parameters. Like prior bias analysis models, stronger parameters of log(2.2) and higher seemed to suggest a lower effect.

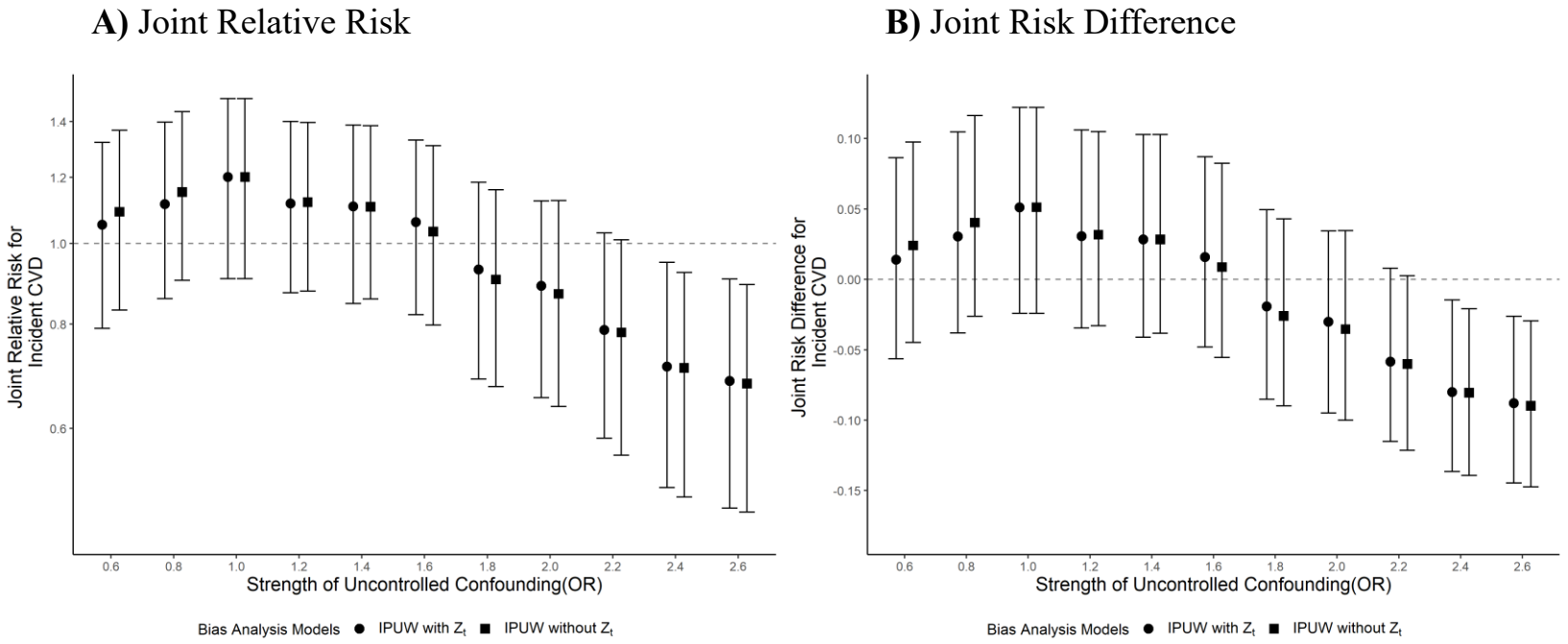
**Figure 4.3. Bias Analysis Where $U_5$ is a Mediator and Confounder for the A) Joint Relative Risk and B) Joint Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants**




Abbreviations: CVD: cardiovascular disease; IPUW: inverse probability of uncontrolled confounder weight; OR: odds ratio; $U_5$: time-varying sedentary behavior at wave 5; $Z_t$: covariate set

Footnote: The bias parameters represent the log(OR) relationship between the exposure, outcome, and uncontrolled confounder at each time point. The value at 1.0 represents the results from the main analysis, i.e., without bias analysis.

**Figure 4.4. Bias Analysis Where $U_5$ is a Confounder only for the A) Joint Relative Risk and B) Joint Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants**




Abbreviations: CVD: cardiovascular disease; IPUW: inverse probability of uncontrolled confounder weight; OR: odds ratio; $U_5$: time-varying sedentary behavior at wave 5; $Z_t$: covariate set

Footnote: The bias parameters represent the log(OR) relationship between the exposure, outcome, and uncontrolled confounder at each time point. The value at 1.0 represents the results from the main analysis, i.e., without bias analysis.

## 4.5 Discussion

We describe and demonstrate the utility of a doubly robust estimation technique to understand and quantify time-varying uncontrolled confounding in combination with g-computation. In analyses, the outcome model was purposely mis-specified in that the $U_t$ were excluded and thus uncontrolled. However, the use of the IPUW for a weighted outcome regression yielded regression parameters that accounted for this uncontrolled confounding. Likewise, misspecification of the IPUW through inclusion or exclusion of covariates further demonstrated the doubly robust properties. We demonstrated this proof of concept in simulations and in our application, we were able to quantify bias due to time-varying uncontrolled confounding using this method. This introduction of a novel bias analysis method addresses the literature gap for bias analysis methods with causal inference for time-varying studies.

Within our applied analyses, we demonstrated a modest effect between persistent elevated depressive symptoms and incident CVD outcomes over a decade of follow-up. Our observations are in the same direction as prior studies that have demonstrated a strong positive relationship between baseline depressive symptoms and CVD outcomes, including incident events and prognosis[26,28,33,40]. Furthermore, HRS analyses from other authors examining a baseline depressive symptoms and CVD outcomes also suggested positive relationships. Using data from the first wave of the HRS in 1992, studies suggested that elevated depressive symptoms were associated with a higher hazard ratio of CVD and odds ratio of heart disease over at least a decade of follow-up[108,109]. These observations were supported by another study of baseline depressive symptoms from the HRS over a four-year follow-up period[90]. To note, the prior studies examined depressive symptoms as a single measurement, had varying measurement tools and/or cutoffs to assess depressive symptoms, and confounder control for lifestyle factors at

baseline and over time remained limited. To address these limitations, studies have recently investigated time-varying effects of depressive symptoms on CVD incidence over shorter two-year periods of follow-up. A study from the China Health and Retirement Longitudinal Study noted that persistent depressive symptoms were associated with higher relative risks of CVD, stroke, and heart disease[38]. Notably, the authors found that persistent symptoms had almost a four times higher risk of stroke compared with no symptoms. Another study of the HRS used a marginal structural model to demonstrate that stable high depressive symptoms were associated with a two-fold higher hazard ratio for stroke compared with stable low or no symptoms over rolling two-year windows[37]. While the focus of this illustration was total CVD rather than individual causes over a longer period of follow-up, we observed a higher risk of CVD with persistent depressive symptoms over a short follow-up period. Our results were imprecise with wide confidence intervals, though were again in the general direction of a previous study. We sought to investigate a decade of follow-up as it aligns with timeframe of the pooled cohort equation that evaluate the CVD risk in CVD-free individuals[79]. Furthermore, our long-term results were somewhat comparable to a recent study of REGARDS (The Reasons for Geographic and Racial Differences in Stroke Study), which also observed a small positive effect between time-varying elevated depressive symptoms and coronary heart disease over a median follow-up of 6.9 years[36]. We too investigated the time-varying relationship of depressive symptoms including estimating the marginal effect estimate with g-computation to adequately account for time-varying confounding. While our observations with somewhat on par with other longitudinal studies and we caution the interpretation given the illustrative purposes, we further sought to understand other biases that may be contributing to the discordance of results.

In our simulations, we demonstrated how an IPUW with a mis-specified g-computation model could result in the true estimate. Thus, we further undertook this bias analysis to quantify time-varying sedentary behavior as the potential uncontrolled confounder and mediator given the noted individual relationships with depressive symptoms and CVD outcomes[41,105,106]. Few studies have distinguished sedentary behavior and physical inactivity, despite several guidelines expanding their recommendations to encompass sedentary behavior[110,111]. A recent study of older women highlighted the linear relationship between total sedentary time with CVD outcomes, independent of physical activity, further supporting the importance in distinction of these two risk factors[106]. In our bias analysis, only strong bias parameters could suggest a protective relationship between persistent elevated depressive symptoms and CVD. Under lower and more plausible parameters, our results were robust even in consideration of sedentary behavior as only a confounder or also a mediator. Given the importance of CVD prevention in the aging population, further studies are warranted to better inform guidelines for depressive symptoms and screening, especially in longitudinal studies.

One of the key features of quantitative bias analysis includes the specification of bias parameters that reflect the magnitude of the relationship with the uncontrolled confounder and other variables. Beyond the implementation of the bias analysis in statistical programming, difficulty could occur when determining the bias parameters, including plausible values. The most common sources in identifying bias parameters for uncontrolled confounders include literature review (external validation data), followed by hypothetical scenarios (or educated guesses) and internal data[49]. Notably, a recent review did not find an application of bias analysis to a time-varying confounder. In our study, we reviewed the literature and used internal data on conceptually related variables (physical inactivity) to better theorize the prevalence of sedentary

behavior and relationships to other variables[41,95,105,106]. Our focus was the adjusted effect estimates relating the uncontrolled confounders, exposures, and outcome to inform our parameters. While each study source may have a differing set of covariates, we demonstrate in simulations that inclusion of covariates are not necessary in the IPUW denominator specification. Our interests are in bias-adjusting the exposures rather than the covariates and thus misspecification of the bias parameters for covariates would ultimately not impact our final goal[97]. Nevertheless, the identification of bias parameters including addressing uncertainty remains an important and careful task given that improper parameters may lead to misleading results[49]. A possible solution to addressing this uncertainty would be a probabilistic bias analysis yet due to computational limitations, we instead performed a series of simple bias analysis models over incrementally increasing parameters. Another key feature presented in this bias analysis is that the role of the unknown variable could pose as either a mediator and/or confounder. Prior studies would suggest sedentary behavior holds a mediating and confounding role, yet often applications of bias analysis would only focus on a single uncontrolled confounder [41,49,105-107,111]. In our analyses, we observed two conclusions in our bias analysis in the presence of strong confounding based on these assumptions. When considered as a confounder only, the bias analysis demonstrated a protective relationship when the parameters were log(2.2) and greater, while this conclusion was less clear when sedentary behavior was also considered as a mediator. It remains important to clarify the role the unknown variable would play, especially in the context of bias analysis and time-varying studies.

There are several limitations to this study. In our simulation 3.3, under more extreme levels of confounding, the IPUW result was slightly underestimated from our true, simulated estimate. Additional work is needed to understand these results. In our models, we chose fixed

bias parameters that were varied over a series of repeated bias analysis models. As a result, these analyses do not incorporate random error and more studies are warranted in understanding how methods of probabilistic bias analysis could include weighting for time-varying uncontrolled confounding. Furthermore, we assumed that we met all other assumptions necessary for g-computation and causal inference. We chose to include all possible two-way interaction terms between our exposures and time-invariant and time-varying covariates. Yet, we acknowledge the possibility of model misspecification including higher order interactions or non-linear relationships in our regressions. While the inclusion of fewer interaction terms yielded similar results to that of our main analyses, the use of advanced modeling techniques would be favorable in more closely identifying model specification. Finally, we only examined one source of bias in uncontrolled confounding and chose to prioritize modifiable lifestyle factors as the main time-varying covariates, given that they are often unavailable in studies. Other potential confounders that were missing or limited in the data include diet, laboratory biomarkers, and detailed prescription data. Bias analysis for other sources of biases including selection bias and measurement error warrant recognition. In the HRS illustration, all data were self-reported and we restricted to HRS participants with two CES-D survey responses during the study period and we acknowledge the possibility of overadjustment coupled with smaller exposure groups. More research is needed in the development of bias analysis for multiple sources of bias.

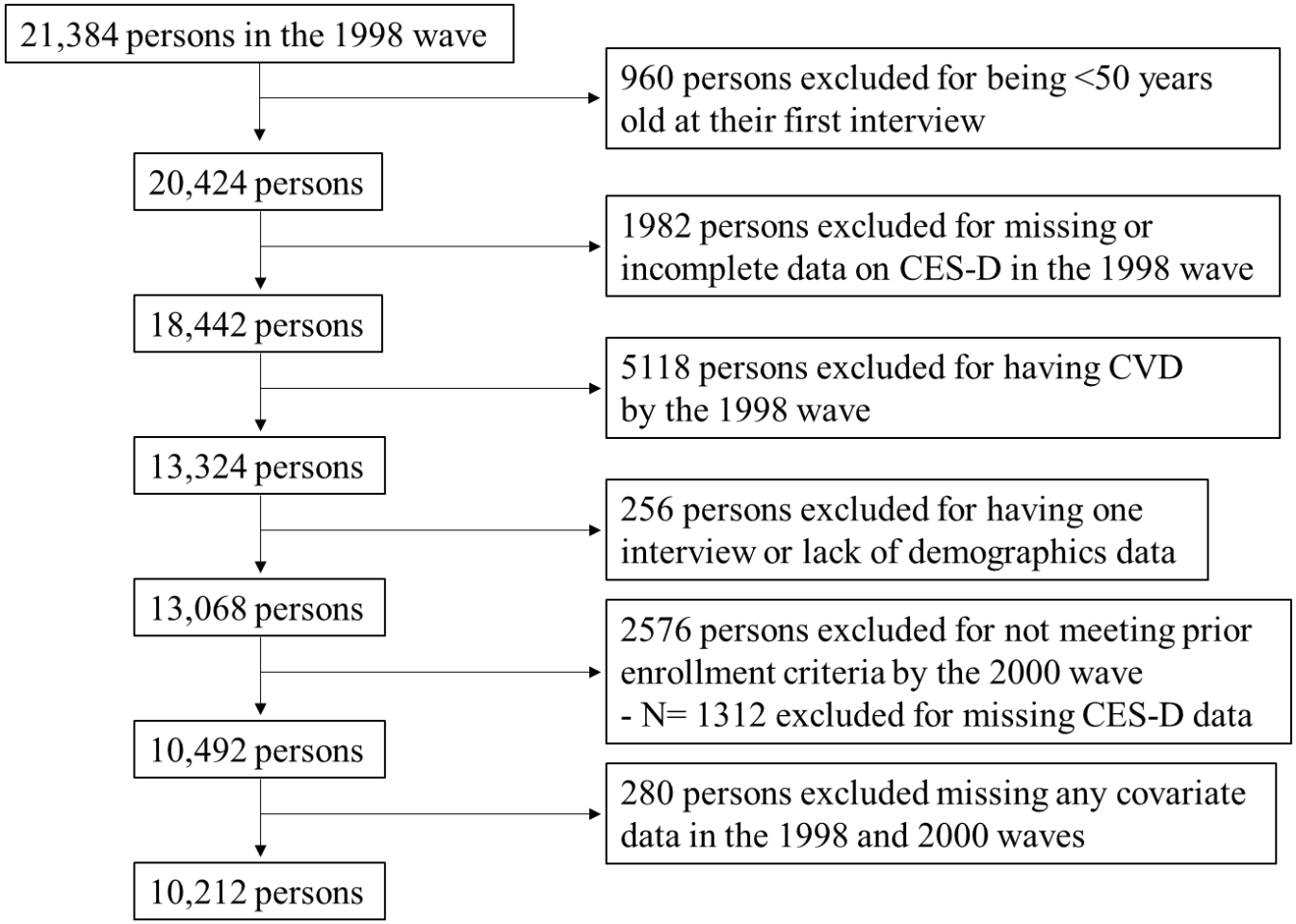
Other strengths include introduction of a novel weighting method in the quantitative bias analysis. Simulations demonstrate the proof of concept to implement this method across different causal structures, effect estimates, and varied parameters. The use of weights has added benefits in that it could be reused without the need for reprogramming or changing parameters due to the analytical data source, and this remains on a records-level. Our simulations and clinical

illustration further demonstrate the utility of the IPUW when the unknown variables could act as only confounders, or also a mediator. We did not use the sampling weights in our HRS analyses due to the time-varying and simulation aspect of our analyses, however the data itself is representative of the older US adult population. The data from the HRS also afforded the ability to investigate repeated measures of key confounders, such as smoking, that are often limited in other longitudinal databases.

In conclusion, this study adds to the bias analysis and causal inference literature by providing a new method for addressing time-varying uncontrolled confounding based on a combined use of weighting and g-computation. We believe that these techniques would be helpful as causal inference using 'big data' from observational studies continue to grow.

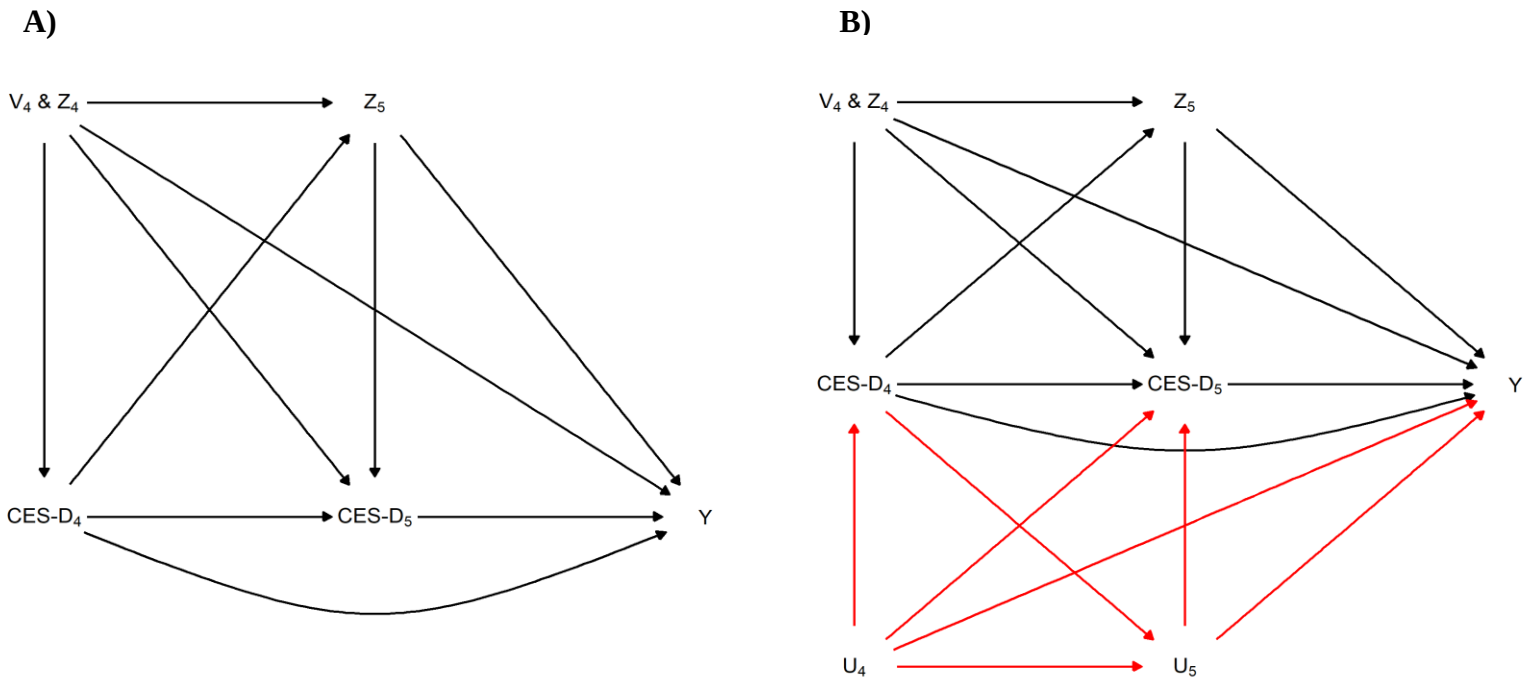
## 4.6 Appendix

**Appendix Figure 4.1. Cohort Construction for Health and Retirement Study Analyses**



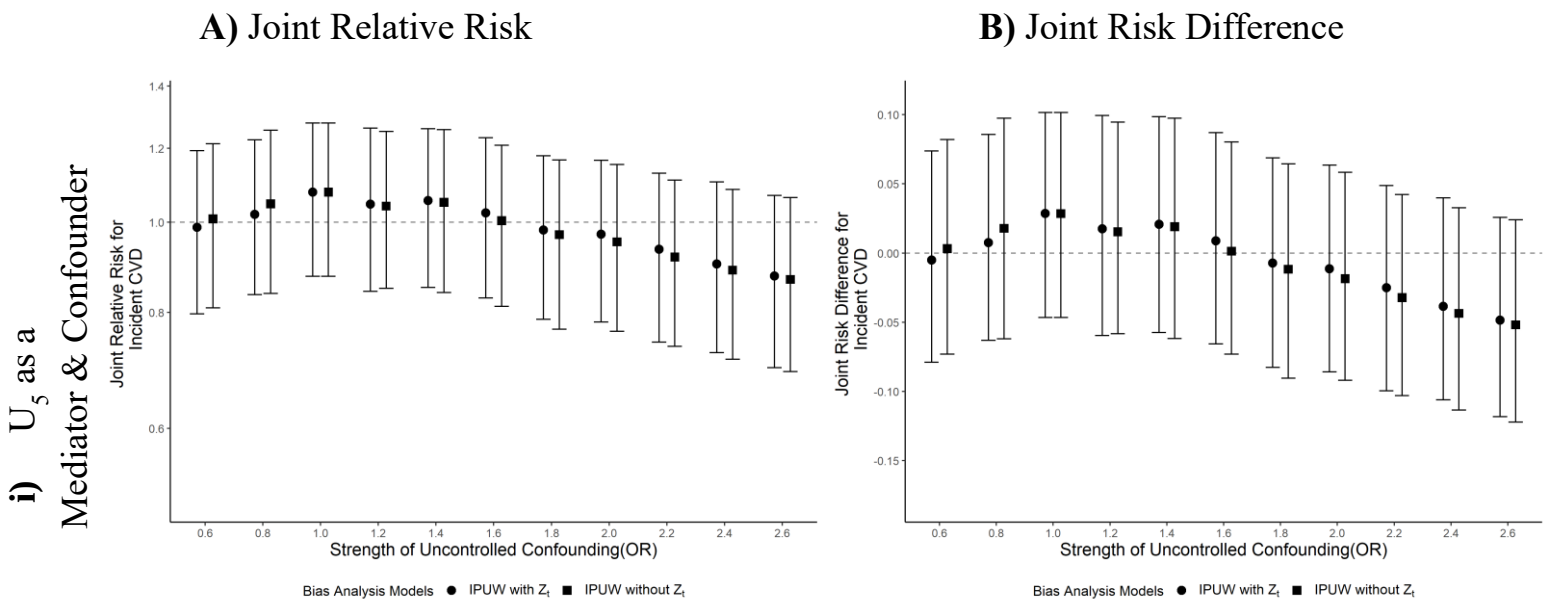

Abbreviations: CES-D: Center for Epidemiological Studies Depression scale score; CVD: cardiovascular disease

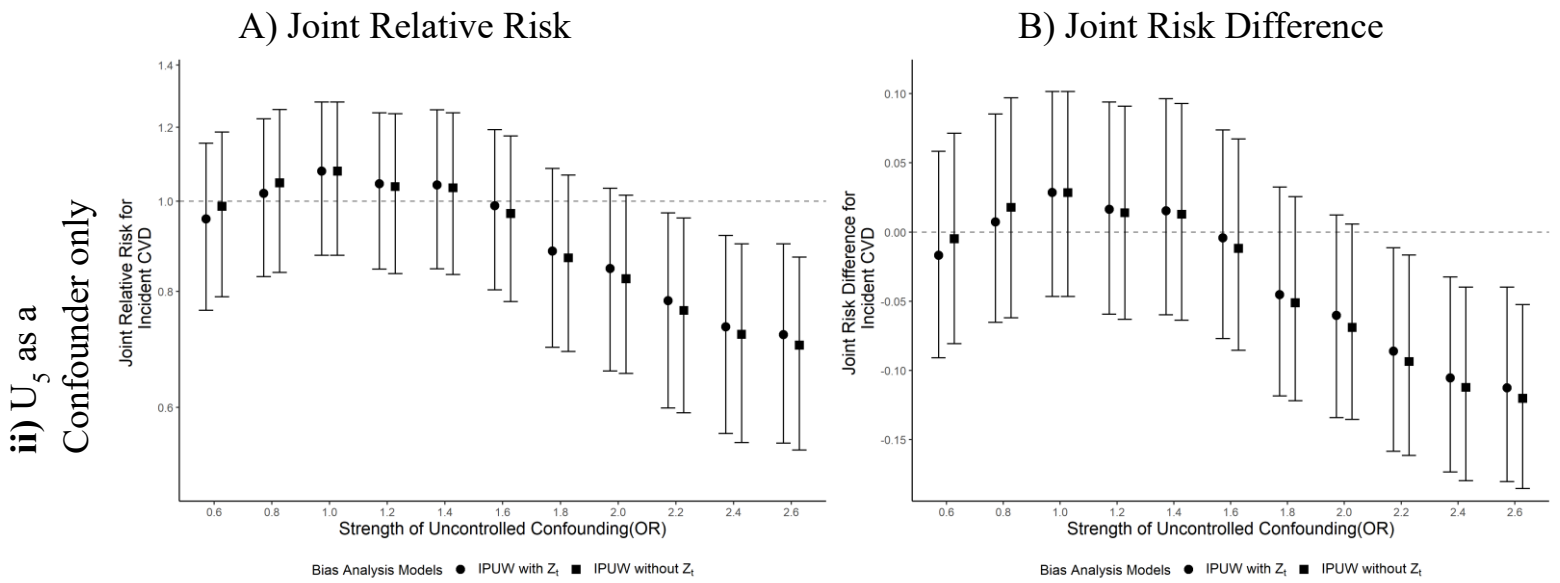
**Appendix Figure 4.2. Directed Acyclic Graphs for the Health and Retirement Study Assessing the Joint Effect of Elevated Depressive Symptoms and Incident CVD A) with and B) without Time-Varying Uncontrolled Confounders**




Abbreviations: CES-D: Center for Epidemiological Studies Depression scale score**;** $V_4$: baseline covariates at wave 4: age, gender, race, ethnicity, high-school equivalent education**;** $Z_4$ and $Z_5$: time-varying covariates at wave 4 and 5, respectively: martial/partnered status, alcohol status, smoking status, body mass index and vigorous physical activity; Y: incident cardiovascular disease after wave 5; $U_4$ and $U_5$: time-varying uncontrolled sedentary behavior at wave 4 and 5, respectively

**Appendix Figure 4.3. Bias Analysis for the Joint A) Relative Risk and B) Risk Difference of Elevated Depressive Symptoms and Incident CVD over Maximum Follow-up in the Health and Retirement Study Participants Where $U_5$ is a i) Mediator and Confounder or ii) a Confounder only**

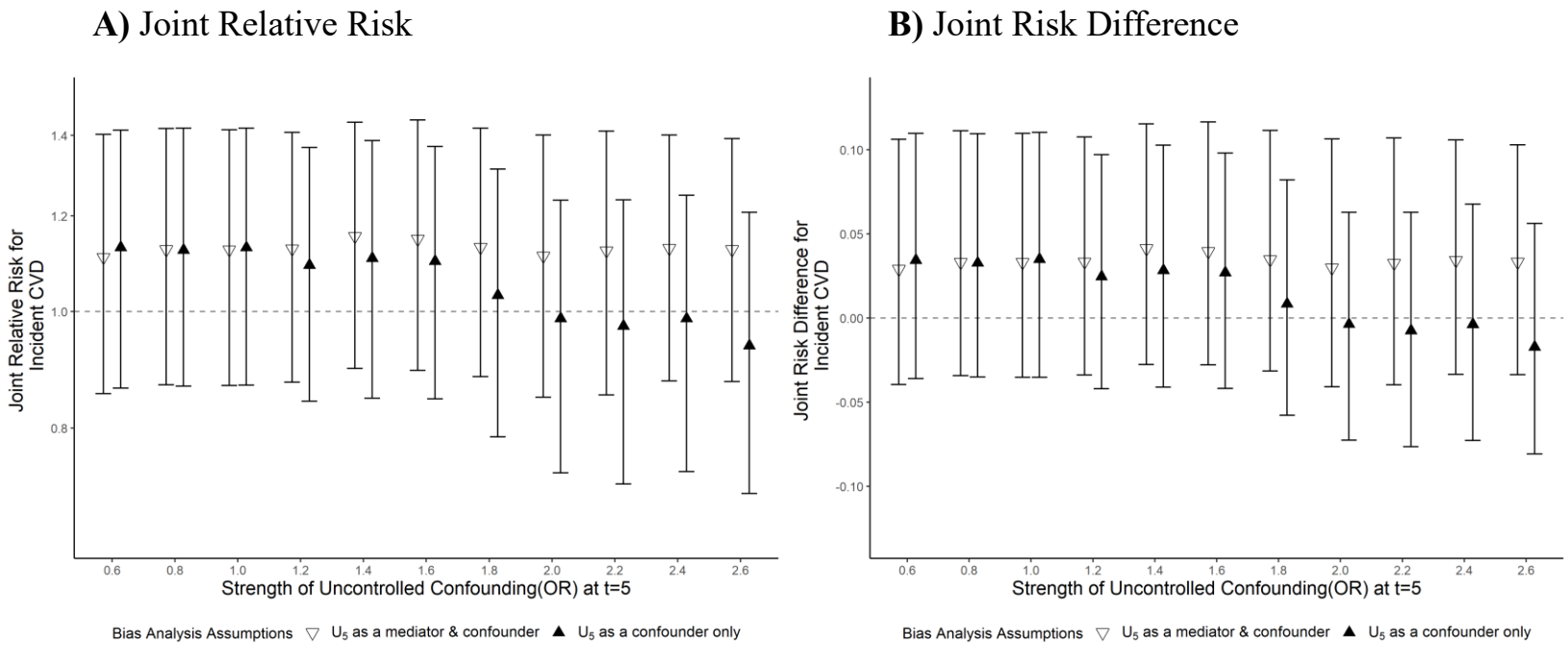





Abbreviations: CVD: cardiovascular disease; IPUW: inverse probability of uncontrolled confounder weight; OR: odds ratio; $U_5$: time-varying sedentary behavior at wave 5; $Z_t$: covariate set

Footnote: The bias parameters represent the log(OR) relationship between the exposure, outcome, and uncontrolled confounder at each time point. The value at 1.0 represents the results from the main analysis, i.e., without bias analysis.

**Appendix Figure 4.4. Bias Analysis for the Joint A) Relative Risk and B) Risk Difference of Persistent Elevated Depressive Symptoms and Incident CVD over Ten Years in the Health and Retirement Study Participants, when Changing the $U_5$ Parameters.**




Abbreviations: CVD: cardiovascular disease; IPUW: inverse probability of uncontrolled confounder weight; OR: odds ratio; t: time point or wave; $U_5$: time-varying sedentary behavior at wave 5

Footnote: The bias parameters represent the log(OR) relationship between the exposure, outcome and uncontrolled confounder at each time point. The value at 1.0 represents the results from the main analysis, i.e., without bias analysis.

**Appendix Table 4.1. Cohort Descriptives of 10,212 Health and Retirement Study Participants Stratified by Elevated Depressive Symptoms Across Waves 4 and 5**

| Characteristic | No elevated depressive symptoms in either Wave 4 or 5 | Elevated depressive symptoms in Wave 4 only | Elevated depressive symptoms in Wave 5 only | Elevated depressive symptoms in both Wave 4 and 5 |
|---|---|---|---|---|
| N (%) | 8237 (80.7%) | 716 (7.0%) | 670 (6.6%) | 589 (5.8%) |
| Age (years) | 63.9 ± 8.8 | 63.4 ± 9.3 | 63.7 ± 10.0 | 64.5 ± 9.3 |
| 50-<60 | 3019 (36.7%) | 294 (41.1%) | 283 (42.2%) | 216 (36.7%) |
| 60-<70 | 3052 (37.1%) | 241 (33.7%) | 191 (28.5%) | 218 (37.0%) |
| 70-<80 | 1713 (20.8%) | 136 (19.0%) | 138 (20.6%) | 98 (16.6%) |
| ≥80 | 453 (5.5%) | 45 (6.3%) | 58 (8.7%) | 57 (9.7%) |
| Gender (% Female) | 4889 (59.4%) | 504 (70.4%) | 505 (75.4%) | 460 (78.1%) |
| Race (%) | | | | |
| White | 7023 (85.3%) | 556 (77.7%) | 538 (80.3%) | 428 (72.7%) |
| African American | 965 (11.7%) | 130 (18.2%) | 109 (16.3%) | 125 (21.2%) |
| Other | 249 (3.0%) | 30 (4.2%) | 23 (3.4%) | 36 (6.1%) |
| Hispanic (%) | 507 (6.2%) | 103 (14.4%) | 68 (10.1%) | 102 (17.3%) |
| Married/Partnered (%) | 6001 (72.9%) | 428 (59.8%) | 454 (67.8%) | 294 (49.9%) |
| Geographic Region (%) | | | | |
| Northeast | 1309 (15.9%) | 128 (17.9%) | 141 (21.1%) | 117 (19.9%) |
| Midwest | 2238 (27.2%) | 143 (20.0%) | 164 (24.5%) | 108 (18.3%) |
| South | 3155 (38.3%) | 310 (43.3%) | 270 (40.4%) | 276 (46.9%) |
| West | 1527 (18.5%) | 135 (18.9%) | 93 (13.9%) | 88 (14.9%) |
| Education (years) | 12.7 ± 3.0 | 11.5 ± 3.5 | 11.5 ± 3.4 | 10.5 ± 3.7 |
| Education ≥12 years (%) | 6570 (79.8%) | 464 (64.8%) | 443 (66.1%) | 310 (52.6%) |
| Lifestyle | | | | |
| Alcohol use (%) | | | | |
| Non-drinker | 5335 (64.8%) | 519 (72.5%) | 489 (73.0%) | 477 (81.0%) |
| Low/Moderate | 1867 (22.7%) | 89 (12.4%) | 92 (13.7%) | 61 (10.4%) |
| Heavy | 1035 (12.6%) | 108 (15.1%) | 89 (13.3%) | 51 (8.7%) |
| Body Mass Index ($kg/m^2$) | 26.9 ± 4.8 | 27.3 ± 5.6 | 27.7 ± 6.1 | 28.1 ± 5.9 |
| <18.5 | 98 (1.2%) | 14 (2.0%) | 11 (1.6%) | 10 (1.7%) |
| 18.5-<25 | 2918 (35.4%) | 242 (33.8%) | 219 (32.7%) | 173 (29.4%) |
| 25-<30 | 3399 (41.3%) | 268 (37.4%) | 253 (37.8%) | 208 (35.3%) |
| 30-<35 | 1332 (16.2%) | 127 (17.7%) | 121 (18.1%) | 131 (22.2%) |
| ≥35 | 490 (5.9%) | 65 (9.1%) | 66 (9.9%) | 67 (11.4%) |
| Current Smoker (%) | 1265 (15.4%) | 152 (21.2%) | 136 (20.3%) | 142 (24.1%) |

| | | | | |
|---|---|---|---|---|
| Physical Activity (%) | 4305 (52.3%) | 309 (43.2%) | 274 (40.9%) | 176 (29.9%) |
| Pre-existing conditions (%) | | | | |
| High Blood Pressure | 3151 (38.3%) | 310 (43.3%) | 300 (44.8%) | 310 (52.7%) |
| Diabetes | 766 (9.3%) | 82 (11.5%) | 98 (14.6%) | 90 (15.3%) |
| Time-updated variables in Wave 5 | | | | |
| Married/Partnered (%) | 5867 (71.2%) | 408 (57.0%) | 391 (58.4%) | 274 (46.5%) |
| Alcohol use (%) | | | | |
| Non-drinker | 5626 (68.3%) | 543 (75.8%) | 516 (77.0%) | 490 (83.2%) |
| Low/Moderate | 1726 (21.0%) | 89 (12.4%) | 86 (12.8%) | 60 (10.2%) |
| Heavy | 885 (10.7%) | 84 (11.7%) | 68 (10.1%) | 39 (6.6%) |
| Body Mass Index ($kg/m^2$) | 27.0 ± 4.9 | 27.5 ± 5.5 | 27.6 ± 6.4 | 28.1 ± 5.9 |
| Current Smoker (%) | 1150 (14.0%) | 151 (21.1%) | 123 (18.4%) | 131 (22.2%) |
| Physical Activity (%) | 4222 (51.3%) | 302 (42.2%) | 232 (34.6%) | 150 (25.5%) |
| Pre-existing conditions (%) | | | | |
| High Blood Pressure | 3441 (41.8%) | 337 (47.1%) | 324 (48.4%) | 339 (57.6%) |
| Diabetes | 891 (10.8%) | 100 (14.0%) | 109 (16.3%) | 113 (19.2%) |

Footnote : Data presented as N(%) or mean±standard deviation, where applicable

## 5. Conclusion

This dissertation set out to develop and demonstrate novel quantitative bias analytical methods for addressing time-varying uncontrolled confounding in epidemiologic studies of time-varying exposures in chronic diseases. Each study sought to emphasize the broader picture and context of why bias analysis remains critical in epidemiology. Electronic medical records and historical population studies with decades of follow-up for events, repeated measurements, and granular data on a variety of previously understudied biomarkers and exposures, are essential sources for research. Causal methodology has afforded researchers the tools to answer complex causal questions, thereby leading to the causal inference revolution. These powerful causal inference studies using big data are increasingly informing public health. As these big data studies grow, their random error concerns diminish but their systematic error concerns grow since they can no longer hide behind random sampling uncertainty. There is, therefore, a critical need for rigorous sensitivity or bias analysis to address systematic error (or bias) and overconfidence in the results. This dissertation adds to the toolkit for quantitative bias analysis and causal inference by addressing time-varying uncontrolled confounding.

Study 1 described the structures of time-varying uncontrolled confounding using directed acyclic graphs. Given the multiple time points of measurements, directed acyclic graphs were critical in illustrating how an uncontrolled confounder would relate to the outcome, next measurement of exposure, next measurement of other covariates and act as a mediator. When using g-computation, the presence of time-varying uncontrolled confounding demonstrated biased effect estimates. Motivation hypothetical examples in cancer epidemiology provided the context within clinical research. Longitudinal studies are affected by time-varying uncontrolled confounding and the magnitude of bias would be exacerbated depending on the timing and

strength of the uncontrolled confounder relationships. Study 1 demonstrated the repercussions of time-varying uncontrolled confounding in research and provided the rationale to pursue sensitivity analysis methods.

Study 2 demonstrated the novel bias analysis method of a bias offset in combination with g-computation. Building on study 1, study 2 showed how a bias offset method using plasmode simulations could remove biasing paths and address the presence of time-varying uncontrolled confounding on effect estimates. Data from the Health and Retirement Study were used to investigate the relationship of joint elevated depressive symptoms and overall, breast, and prostate cancer incidence in a cohort of older adults. This analysis observed a very small and almost null joint risk difference for each outcome, although challenges in precision made it difficult for estimation. This study failed to find a relationship between elevated depressive symptoms and overall cancer incidence and were of a similar conclusion to other studies. Subsequently, the use of a bias offset to adjust for unmeasured time-varying sedentary behavior suggested that the results were generally robust. Within the context of cancer epidemiology, additional studies are needed to better understand the relationship of longitudinal elevated depressive symptoms and cancer incidence, with a specific focus on site-specific cancers. Within the context of bias analysis, study 2 showcased the application of a bias offset as one method in addressing time-varying uncontrolled confounding.

Study 3 examined the combination of an inverse probability of uncontrolled confounding weight and g-computation to examine time-varying uncontrolled confounding. Akin to study 2, study 3 demonstrated the development of an inverse probability of uncontrolled confounding weight used with a mis-specified g-computation outcome model to obtain a bias-adjusted estimate. Using data from the Health and Retirement Study, study 3 investigated the relationship

of persistent elevated depressive symptoms with incidence of cardiovascular disease, compared with lower depressive symptoms at all times. The g-computation results suggested a modest joint relative risk and risk difference that was consistent across a series of sensitivity analyses, albeit with some difficulties in precision. Study 3 also examined time-varying sedentary behavior in bias analysis. It showed that the magnitude of the relationships with time-varying uncontrolled confounding as well as its role (i.e., a confounder only or a confounder and mediator), would impact results. Thus, study 3 illustrated another bias analysis method with g-computation and called for caution when determining the bias parameters and causal structure.

This dissertation described time-varying uncontrolled confounding and introduced novel bias analysis methods in combination with g-computation. Its illustrations using the Health and Retirement Study exemplify how longitudinal and covariate-rich datasets could inform public health and results should be subject to bias analysis. In addition to the public health implications and importance of sensitivity analyses, this dissertation could also inform additional work in bias analysis. Future avenues in the context of time-varying uncontrolled confounders include extension of these methods to encompass other distributions and forms of the uncontrolled confounders, utilize other methods and effect estimates featured in causal inference such as g-estimation or inverse probability weighting and pooled or cumulative effects, respectively, and applications to studies of more time points of follow-up to address the more complex time-varying structures. Furthermore, this work only addressed one source of bias. Future research should investigate methods to address multiple sources of biases. Although epidemiology has been at the forefront of causal inference and bias analysis, it has focused largely on single time point or point exposure studies in addressing uncontrolled confounding. This dissertation is an ambitious attempt to close the gap by extending and applying bias modeling to time-varying

uncontrolled confounding. Hopefully, this work will stimulate more frequent use of quantitative bias analysis to evaluate the sensitivity of the results of empirical observational studies to uncontrolled confounding that changes as the exposures change.

**References**


1. Benjamin EJ, Muntner P, Alonso A, et al. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. *Circulation*. 03 2019;139(10):e56-e528. doi:10.1161/CIR.0000000000000659

2. Mokdad AH, Ballestros K, Echko M, et al. The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. *JAMA*. 04 2018;319(14):1444-1472. doi:10.1001/jama.2018.0158

3. Waters H, Graf M. *The Cost of Chronic Diseases in the US*. 2018. https://milkeninstitute.org/sites/default/files/reports-pdf/ChronicDiseases-HighRes-FINAL_0.pdf

4. Papanicolas I, Woskie LR, Jha AK. Health Care Spending in the United States and Other High-Income Countries. *JAMA*. Mar 2018;319(10):1024-1039. doi:10.1001/jama.2018.1150

5. Lichtman JH, Froelicher ES, Blumenthal JA, et al. Depression as a risk factor for poor prognosis among patients with acute coronary syndrome: systematic review and recommendations: a scientific statement from the American Heart Association. *Circulation*. Mar 2014;129(12):1350-69. doi:10.1161/CIR.0000000000000019

6. Lash TL, Fox MP, Cooney D, Lu Y, Forshee RA. Quantitative Bias Analysis in Regulatory Settings. *Am J Public Health*. 07 2016;106(7):1227-30. doi:10.2105/AJPH.2016.303199

7. Murray CJ, Atkinson C, Bhalla K, et al. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. *JAMA*. Aug 2013;310(6):591-608. doi:10.1001/jama.2013.13805

8. Villarroel M, Terlizzi E. Symptoms of depression among adults: United States, 2019. NCHS Data Brief, no 379. Hyattsville, MD: National Center for Health Statistics. 2020.

9. Kocarnik JM, Compton K, Dean FE, et al. Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life Years for 29 Cancer Groups From 2010 to 2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *JAMA Oncol*. 03 01 2022;8(3):420-444. doi:10.1001/jamaoncol.2021.6987

10. Gross AL, Gallo JJ, Eaton WW. Depression and cancer risk: 24 years of follow-up of the Baltimore Epidemiologic Catchment Area sample. *Cancer Causes Control*. Feb 2010;21(2):191-9. doi:10.1007/s10552-009-9449-1

11. Jia Y, Li F, Liu YF, Zhao JP, Leng MM, Chen L. Depression and cancer risk: a systematic review and meta-analysis. *Public Health*. Aug 2017;149:138-148. doi:10.1016/j.puhe.2017.04.026

12. Chang HY, Keyes KM, Mok Y, Jung KJ, Shin YJ, Jee SH. Depression as a risk factor for overall and hormone-related cancer: the Korean cancer prevention study. *J Affect Disord*. Mar 2015;173:1-8. doi:10.1016/j.jad.2014.10.064

13. Dalton SO, Mellemkjaer L, Olsen JH, Mortensen PB, Johansen C. Depression and cancer risk: a register-based study of patients hospitalized with affective disorders, Denmark, 1969-1993. *Am J Epidemiol*. Jun 2002;155(12):1088-95. doi:10.1093/aje/155.12.1088

14. Aro AR, De Koning HJ, Schreck M, Henriksson M, Anttila A, Pukkala E. Psychological risk factors of incidence of breast cancer: a prospective cohort study in Finland. *Psychol Med*. Oct 2005;35(10):1515-21. doi:10.1017/S0033291705005313

15. Oerlemans ME, van den Akker M, Schuurman AG, Kellen E, Buntinx F. A meta-analysis on depression and subsequent cancer risk. *Clin Pract Epidemiol Ment Health*. Dec 2007;3:29. doi:10.1186/1745-0179-3-29

16. Sun HL, Dong XX, Cong YJ, et al. Depression and the risk of breast cancer: a meta-analysis of cohort studies. *Asian Pac J Cancer Prev*. 2015;16(8):3233-9. doi:10.7314/apjcp.2015.16.8.3233

17. Wang YH, Li JQ, Shi JF, et al. Depression and anxiety in relation to cancer incidence and mortality: a systematic review and meta-analysis of cohort studies. *Mol Psychiatry*. 07 2020;25(7):1487-1499. doi:10.1038/s41380-019-0595-x

18. Chen X, Kong J, Diao X, et al. Depression and prostate cancer risk: A Mendelian randomization study. *Cancer Med*. Oct 2020;doi:10.1002/cam4.3493

19. Huang T, Poole EM, Okereke OI, et al. Depression and risk of epithelial ovarian cancer: Results from two large prospective cohort studies. *Gynecol Oncol*. Dec 2015;139(3):481-6. doi:10.1016/j.ygyno.2015.10.004

20. van Tuijl LA, Voogd AC, de Graeff A, et al. Psychosocial factors and cancer incidence (PSY-CA): Protocol for individual participant data meta-analyses. *Brain Behav*. 10 2021;11(10):e2340. doi:10.1002/brb3.2340

21. Archer G, Pikhart H, Head J. Do depressive symptoms predict cancer incidence?: 17-year follow-up of the Whitehall II study. *J Psychosom Res*. Dec 2015;79(6):595-603. doi:10.1016/j.jpsychores.2015.07.011

22. Lemogne C, Consoli SM, Melchior M, et al. Depression and the risk of cancer: a 15-year follow-up study of the GAZEL cohort. *Am J Epidemiol*. Dec 15 2013;178(12):1712-20. doi:10.1093/aje/kwt217

23. Roth GA, Mensah GA, Johnson CO, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *J Am Coll Cardiol*. 12 22 2020;76(25):2982-3021. doi:10.1016/j.jacc.2020.11.010

24. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. *Circulation*. 02 22 2022;145(8):e153-e639. doi:10.1161/CIR.0000000000001052

25. Templin C, Ghadri JR, Diekmann J, et al. Clinical Features and Outcomes of Takotsubo (Stress) Cardiomyopathy. *N Engl J Med*. Sep 2015;373(10):929-38. doi:10.1056/NEJMoa1406761

26. Rutledge T, Gould H. *Chapter 2 - Epidemiological evidence linking stress and depression with CVD*. 2020:15 - 34.

27. Surtees PG, Wainwright NW, Luben RN, Wareham NJ, Bingham SA, Khaw KT. Depression and ischemic heart disease mortality: evidence from the EPIC-Norfolk United Kingdom prospective cohort study. *Am J Psychiatry*. Apr 2008;165(4):515-23. doi:10.1176/appi.ajp.2007.07061018

28. Harshfield EL, Pennells L, Schwartz JE, et al. Association Between Depressive Symptoms and Incident Cardiovascular Diseases. *JAMA*. 12 15 2020;324(23):2396-2405. doi:10.1001/jama.2020.23068

29. Huang CJ, Hsieh MH, Hou WH, Liu JC, Jeng C, Tsai PS. Depression, antidepressants, and the risk of coronary heart disease: a population-based cohort study. *Int J Cardiol*. Oct 12 2013;168(5):4711-6. doi:10.1016/j.ijcard.2013.07.173

30. Eurelings LS, van Dalen JW, Ter Riet G, et al. Apathy and depressive symptoms in older people and incident myocardial infarction, stroke, and mortality: a systematic review and meta-analysis of individual participant data. *Clin Epidemiol*. 2018;10:363-379. doi:10.2147/CLEP.S150915

31. Correll CU, Solmi M, Veronese N, et al. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry*. Jun 2017;16(2):163-180. doi:10.1002/wps.20420

32. Van der Kooy K, van Hout H, Marwijk H, Marten H, Stehouwer C, Beekman A. Depression and the risk for cardiovascular diseases: systematic review and meta analysis. *Int J Geriatr Psychiatry*. Jul 2007;22(7):613-26. doi:10.1002/gps.1723

33. Gan Y, Gong Y, Tong X, et al. Depression and the risk of coronary heart disease: a meta-analysis of prospective cohort studies. *BMC Psychiatry*. Dec 2014;14:371. doi:10.1186/s12888-014-0371-z

34. Larkin K, Chantler P. *Chapter 1 - Stress, depression, and cardiovascular disease*. 2020:1 - 12.

35. Whiteford HA, Harris MG, McKeon G, et al. Estimating remission from untreated major depression: a systematic review and meta-analysis. *Psychol Med*. Aug 2013;43(8):1569-85. doi:10.1017/S0033291712001717

36. Moise N, Khodneva Y, Richman J, Shimbo D, Kronish I, Safford MM. Elucidating the Association Between Depressive Symptoms, Coronary Heart Disease, and Stroke in Black and White Adults: The REasons for Geographic And Racial Differences in Stroke (REGARDS) Study. *J Am Heart Assoc*. 08 2016;5(8)doi:10.1161/JAHA.116.003767

37. Gilsanz P, Walter S, Tchetgen Tchetgen EJ, et al. Changes in Depressive Symptoms and Incidence of First Stroke Among Middle-Aged and Older US Adults. *J Am Heart Assoc*. May 13 2015;4(5)doi:10.1161/JAHA.115.001923

38. Li H, Qian F, Hou C, et al. Longitudinal Changes in Depressive Symptoms and Risks of Cardiovascular Disease and All-Cause Mortality: A Nationwide Population-Based Cohort Study. *J Gerontol A Biol Sci Med Sci*. 10 15 2020;75(11):2200-2206. doi:10.1093/gerona/glz228

39. Everson-Rose SA, Roetker NS, Lutsey PL, et al. Chronic stress, depressive symptoms, anger, hostility, and risk of stroke and transient ischemic attack in the multi-ethnic study of atherosclerosis. *Stroke*. Aug 2014;45(8):2318-23. doi:10.1161/STROKEAHA.114.004815

40. Nicholson A, Kuper H, Hemingway H. Depression as an aetiologic and prognostic factor in coronary heart disease: a meta-analysis of 6362 events among 146 538 participants in 54 observational studies. *Eur Heart J*. Dec 2006;27(23):2763-74. doi:10.1093/eurheartj/ehl338

41. Brummett BH, Babyak MA, Siegler IC, Mark DB, Williams RB, Barefoot JC. Effect of smoking and sedentary behavior on the association between depressive symptoms and mortality from coronary heart disease. *Am J Cardiol*. Sep 01 2003;92(5):529-32. doi:10.1016/s0002-9149(03)00719-7

42. Ziegelstein RC. Improving depression and reducing cardiac events: which is the chicken and which is the egg? *J Psychosom Res*. May 2013;74(5):454-7. doi:10.1016/j.jpsychores.2013.02.005

43. Rothman K, Greenland S, Lash T. *Modern Epidemiology, 3rd Edition*. Lippincott Williams & Wilkins; 2008.

44. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. Sep 2000;11(5):550-60. doi:10.1097/00001648-200009000-00011

45. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med*. Apr 2013;32(9):1584-618. doi:10.1002/sim.5686

46. Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding-a systematic review of the literature. *Int J Epidemiol*. 02 2019;48(1):254-265. doi:10.1093/ije/dyy218

47. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. Dec 2014;43(6):1969-85. doi:10.1093/ije/dyu149

48. Arah OA, Chiba Y, Greenland S. Bias formulas for external adjustment and sensitivity analysis of unmeasured confounders. *Ann Epidemiol*. Aug 2008;18(8):637-46. doi:10.1016/j.annepidem.2008.04.003

49. Petersen JM, Ranker LR, Barnard-Mayers R, MacLehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. *Int J Epidemiol*. 11 10 2021;50(5):1708-1730. doi:10.1093/ije/dyab061

50. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology*. Jan 2011;22(1):42-52. doi:10.1097/EDE.0b013e3181f74493

51. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. Aug 2017;167(4):268-274. doi:10.7326/M16-2607

52. Arah OA. Bias Analysis for Uncontrolled Confounding in the Health Sciences. *Annu Rev Public Health*. Mar 2017;38:23-38. doi:10.1146/annurev-publhealth-032315-021644

53. Amoon AT, Arah OA, Kheifets L. The sensitivity of reported effects of EMF on childhood leukemia to uncontrolled confounding by residential mobility: a hybrid simulation study and an empirical analysis using CAPS data. *Cancer Causes Control*. Aug 2019;30(8):901-908. doi:10.1007/s10552-019-01189-9

54. Thompson CA, Zhang ZF, Arah OA. Competing risk bias to explain the inverse relationship between smoking and malignant melanoma. *Eur J Epidemiol*. Jul 2013;28(7):557-67. doi:10.1007/s10654-013-9812-0

55. Lin SH, Young J, Logan R, Tchetgen Tchetgen EJ, VanderWeele TJ. Parametric Mediational g-Formula Approach to Mediation Analysis with Time-varying Exposures,

Mediators, and Confounders. *Epidemiology*. 03 2017;28(2):266-274. doi:10.1097/EDE.0000000000000609

56. Brumback BA, Hernán MA, Haneuse SJ, Robins JM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Stat Med*. Mar 2004;23(5):749-67. doi:10.1002/sim.1657

57. Burne RM, Abrahamowicz M. Adjustment for time-dependent unmeasured confounders in marginal structural Cox models using validation sample data. *Stat Methods Med Res*. 02 2019;28(2):357-371. doi:10.1177/0962280217726800

58. Cole SR, Hernán MA, Margolick JB, Cohen MH, Robins JM. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. *Am J Epidemiol*. Sep 2005;162(5):471-8. doi:10.1093/aje/kwi216

59. Soohoo M, Obi Y, Rivara MB, et al. Comparative Effectiveness of Dialysis Modality on Laboratory Parameters of Mineral Metabolism. *Am J Nephrol*. Feb 28 2022:1-12. doi:10.1159/000521508

60. Yang S, Eaton CB, Lu J, Lapane KL. Application of marginal structural models in pharmacoepidemiologic studies: a systematic review. *Pharmacoepidemiol Drug Saf*. Jun 2014;23(6):560-71. doi:10.1002/pds.3569

61. Klungsøyr O, Sexton J, Sandanger I, Nygård JF. Sensitivity analysis for unmeasured confounding in a marginal structural Cox proportional hazards model. *Lifetime Data Anal*. Jun 2009;15(2):278-94. doi:10.1007/s10985-008-9109-x

62. Shen C, Li X, Li L, Were MC. Sensitivity analysis for causal inference using inverse probability weighting. *Biom J*. Sep 2011;53(5):822-37. doi:10.1002/bimj.201100042

63. Blackwell M, Yamauchi S. Adjusting for Unmeasured Confounding in Marginal Structural Models with Propensity-Score Fixed Effects. 2021

64. Pearl J. *Causality*. Cambridge: Cambridge University Press.; 2009.

65. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. Apr 2011;173(7):731-8. doi:10.1093/aje/kwq472

66. Calip GS, Boudreau DM, Loggers ET. Changes in adherence to statins and subsequent lipid profiles during and following breast cancer treatment. *Breast Cancer Res Treat*. Feb 2013;138(1):225-33. doi:10.1007/s10549-013-2424-2

67. Jensen BW, Bjerregaard LG, Ängquist L, et al. Change in weight status from childhood to early adulthood and late adulthood risk of colon cancer in men: a population-

based cohort study. *Int J Obes (Lond)*. 10 2018;42(10):1797-1803. doi:10.1038/s41366-018-0109-y

68. Jensen BW, Meyle KD, Madsen K, Sørensen TIA, Baker JL. Early life body size in relation to risk of renal cell carcinoma in adulthood: a Danish observational cohort study. *Eur J Epidemiol*. Mar 2020;35(3):251-258. doi:10.1007/s10654-020-00605-8

69. Carsey TM, Harden JJ. *Monte Carlo simulation and resampling methods for social science .* 2014.

70. Mehta LS, Watson KE, Barac A, et al. Cardiovascular Disease and Breast Cancer: Where These Entities Intersect: A Scientific Statement From the American Heart Association. *Circulation*. 02 2018;137(8):e30-e66. doi:10.1161/CIR.0000000000000556

71. Manthravadi S, Shrestha A, Madhusudhana S. Impact of statin use on cancer recurrence and mortality in breast cancer: A systematic review and meta-analysis. *Int J Cancer*. Sep 15 2016;139(6):1281-8. doi:10.1002/ijc.30185

72. Lauffenburger JC, Gagne JJ, Song Z, Brill G, Choudhry NK. Potentially disruptive life events: what are the immediate impacts on chronic disease management? A case-crossover analysis. *BMJ Open*. 08 2016;6(8):e010958. doi:10.1136/bmjopen-2015-010958

73. Mann DM, Woodward M, Muntner P, Falzon L, Kronish I. Predictors of nonadherence to statins: a systematic review and meta-analysis. *Ann Pharmacother*. Sep 2010;44(9):1410-21. doi:10.1345/aph.1P150

74. Navar AM, Roe MT, White JA, et al. Medication Discontinuation in the IMPROVE-IT Trial. *Circ Cardiovasc Qual Outcomes*. 01 2019;12(1):e005041. doi:10.1161/CIRCOUTCOMES.118.005041

75. Lansberg P, Lee A, Lee ZV, Subramaniam K, Setia S. Nonadherence to statins: individualized intervention strategies outside the pill box. *Vasc Health Risk Manag*. 2018;14:91-102. doi:10.2147/VHRM.S158641

76. Aarestrup J, Bjerregaard LG, Meyle KD, et al. Birthweight, childhood overweight, height and growth and adult cancer risks: a review of studies using the Copenhagen School Health Records Register. *Int J Obes (Lond)*. Jul 2020;44(7):1546-1560. doi:10.1038/s41366-020-0523-9

77. Zohar L, Rottenberg Y, Twig G, et al. Adolescent overweight and obesity and the risk for pancreatic cancer among men and women: a nationwide study of 1.79 million Israeli adolescents. *Cancer*. 01 2019;125(1):118-126. doi:10.1002/cncr.31764

78. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*. May 2014;25(3):418-26. doi:10.1097/EDE.0000000000000078

79. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. Jul 2014;63(25 Pt B):2935-2959. doi:10.1016/j.jacc.2013.11.005

80. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol*. 06 2019;73(24):e285-e350. doi:10.1016/j.jacc.2018.11.003

81. The NS, Suchindran C, North KE, Popkin BM, Gordon-Larsen P. Association of adolescent obesity with risk of severe obesity in adulthood. *JAMA*. Nov 2010;304(18):2042-7. doi:10.1001/jama.2010.1635

82. Emmett PM, Jones LR. Diet, growth, and obesity development throughout childhood in the Avon Longitudinal Study of Parents and Children. *Nutr Rev*. Oct 2015;73 Suppl 3:175-206. doi:10.1093/nutrit/nuv054

83. Chiolero A, Paradis G, Kaufman JS. Assessing the possible direct effect of birth weight on childhood blood pressure: a sensitivity analysis. *Am J Epidemiol*. Jan 01 2014;179(1):4-11. doi:10.1093/aje/kwt228

84. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort Profile: the Health and Retirement Study (HRS). *Int J Epidemiol*. Apr 2014;43(2):576-85. doi:10.1093/ije/dyu067

85. Data from: RAND HRS Longitudinal File 2018 (V1). *Santa Monica, CA*. Deposited Feb 2021.

86. Steffick D, Wallace R, Herzog A, Ofstedal M, Fonda S, Langa K. Documentation of Affective Functioning Measures in the Health and Retirement Study. Ann Arbor, MI2000.

87. Marden JR, Tchetgen Tchetgen EJ, Kawachi I, Glymour MM. Contribution of Socioeconomic Status at 3 Life-Course Periods to Late-Life Memory Function and Decline: Early and Late Predictors of Dementia Risk. *Am J Epidemiol*. Oct 01 2017;186(7):805-814. doi:10.1093/aje/kwx155

88. Dang L, Dong L, Mezuk B. Shades of Blue and Gray: A Comparison of the Center for Epidemiologic Studies Depression Scale and the Composite International Diagnostic

Interview for Assessment of Depression Syndrome in Later Life. *Gerontologist*. 05 15 2020;60(4):e242-e253. doi:10.1093/geront/gnz044

89. *Health and Retirement Study, Cancer Site (v7.0) restricted use dataset*. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740) Ann Arbor, MI. 2022.

90. Niles AN, O'Donovan A. Comparing anxiety and depression to obesity and smoking as predictors of major medical illnesses and somatic symptoms. *Health Psychol*. Feb 2019;38(2):172-181. doi:10.1037/hea0000707

91. Kerr J, Anderson C, Lippman SM. Physical activity, sedentary behaviour, diet, and cancer: an update and emerging new evidence. *Lancet Oncol*. 08 2017;18(8):e457-e471. doi:10.1016/S1470-2045(17)30411-4

92. Teras LR, Gapstur SM, Diver WR, Birmann BM, Patel AV. Recreational physical activity, leisure sitting time and risk of non-Hodgkin lymphoid neoplasms in the American Cancer Society Cancer Prevention Study II Cohort. *Int J Cancer*. Oct 15 2012;131(8):1912-20. doi:10.1002/ijc.27445

93. Zhai L, Zhang Y, Zhang D. Sedentary behaviour and the risk of depression: a meta-analysis. *Br J Sports Med*. Jun 2015;49(11):705-9. doi:10.1136/bjsports-2014-093613

94. Biswas A, Oh PI, Faulkner GE, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med*. Jan 20 2015;162(2):123-32. doi:10.7326/M14-1651

95. Matthews CE, George SM, Moore SC, et al. Amount of time spent in sedentary behaviors and cause-specific mortality in US adults. *Am J Clin Nutr*. Feb 2012;95(2):437-45. doi:10.3945/ajcn.111.019620

96. Patel AV, Maliniak ML, Rees-Punia E, Matthews CE, Gapstur SM. Prolonged Leisure Time Spent Sitting in Relation to Cause-Specific Mortality in a Large US Cohort. *Am J Epidemiol*. 10 01 2018;187(10):2151-2158. doi:10.1093/aje/kwy125

97. Soohoo M, Streja E, Hsiung JT, Kovesdy CP, Kalantar-Zadeh K, Arah OA. Cohort Study and Bias Analysis of the Obesity Paradox Across Stages of Chronic Kidney Disease. *J Ren Nutr*. Nov 30 2021;doi:10.1053/j.jrn.2021.10.007

98. Thompson CA, Arah OA. Selection bias modeling using observed data augmented with imputed record-level probabilities. *Ann Epidemiol*. Oct 2014;24(10):747-53. doi:10.1016/j.annepidem.2014.07.014

99. Vansteelandt S, Keiding N. Invited commentary: G-computation--lost in translation? *Am J Epidemiol*. Apr 01 2011;173(7):739-42. doi:10.1093/aje/kwq474

100. Arah OA, Sudan M, Olsen J, Kheifets L. Marginal structural models, doubly robust estimation, and bias analysis in perinatal and paediatric epidemiology. *Paediatr Perinat Epidemiol*. May 2013;27(3):263-5. doi:10.1111/ppe.12049

101. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. Jan 2009;20(1):18-26. doi:10.1097/EDE.0b013e31818f69ce

102. Yu Y, Liew Z, Wang A, et al. Mediating roles of preterm birth and restricted fetal growth in the relationship between maternal education and infant mortality: A Danish population-based cohort study. *PLoS Med*. 06 2019;16(6):e1002831. doi:10.1371/journal.pmed.1002831

103. Whooley MA, Wong JM. Depression and cardiovascular disorders. *Annu Rev Clin Psychol*. 2013;9:327-54. doi:10.1146/annurev-clinpsy-050212-185526

104. Inoue K, Mayeda ER, Nianogo R, et al. Estimating the joint effect of diabetes and subsequent depressive symptoms on mortality among older latinos. *Ann Epidemiol*. 12 2021;64:120-126. doi:10.1016/j.annepidem.2021.09.007

105. Huang Y, Li L, Gan Y, et al. Sedentary behaviors and risk of depression: a meta-analysis of prospective studies. *Transl Psychiatry*. 01 22 2020;10(1):26. doi:10.1038/s41398-020-0715-z

106. Bellettiere J, LaMonte MJ, Evenson KR, et al. Sedentary behavior and cardiovascular disease in older women: The Objective Physical Activity and Cardiovascular Health (OPACH) Study. *Circulation*. 02 19 2019;139(8):1036-1046. doi:10.1161/CIRCULATIONAHA.118.035312

107. Stubbs B, Vancampfort D, Firth J, et al. Relationship between sedentary behavior and depression: A mediation analysis of influential factors across the lifespan among 42,469 people in low- and middle-income countries. *J Affect Disord*. 03 15 2018;229:231-238. doi:10.1016/j.jad.2017.12.104

108. Xiang X, An R. Depression and onset of cardiovascular disease in the US middle-aged and older adults. *Aging Ment Health*. 2015;19(12):1084-92. doi:10.1080/13607863.2014.1003281

109. Karakus MC, Patton LC. Depression and the onset of chronic illness in older adults: a 12-year prospective study. *J Behav Health Serv Res*. Jul 2011;38(3):373-82. doi:10.1007/s11414-011-9234-2

110. Thivel D, Tremblay A, Genin PM, Panahi S, Rivière D, Duclos M. Physical Activity, Inactivity, and Sedentary Behaviors: Definitions and Implications in Occupational Health. *Front Public Health*. 2018;6:288. doi:10.3389/fpubh.2018.00288

111. Duran AT, Romero E, Diaz KM. Is Sedentary Behavior a Novel Risk Factor for Cardiovascular Disease? *Curr Cardiol Rep*. Feb 19 2022;doi:10.1007/s11886-022-01657-w