FORECASTING THE BURDEN OF DEMENTIA IN THE UNITED STATES

Jordan Weiss

A DISSERTATION

in

Demography and Sociology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

_____

Michel Guillot
Professor of Sociology

Graduate Group Chairperson, Demography

_____

Hans-Peter Kohler, Frederick J. Warren Professor of Demography

Graduate Group Chairperson, Sociology

_____

Jason Schnittker, Professor of Sociology

Dissertation Committee
Michel Guillot, Professor of Sociology
Samuel H. Preston, Professor of Sociology
Norma B. Coe, Associate Professor of Medical Ethics and Health Policy

FORECASTING THE BURDEN OF DEMENTIA IN THE UNITED STATES

ACKNOWLEDGEMENTS

This work was supported by NIH grant T32 HD007242 awarded to the Population Studies Center at the University of Pennsylvania and by NIH grants R01 AG049815 and R01 AG057501 awarded to Norma Coe. I am grateful to Michel Guillot, Sam Preston, and Norma Coe for their valuable guidance and support. I would also like to thank Irma Elo, Doug Ewbank, Jason Karlawish, Jason Schnittker, and Dylan Small for our inspiring conversations spanning population aging, biodemography, cognitive health, medical sociology, and causal inference.

ABSTRACT

FORECASTING THE BURDEN OF DEMENTIA IN THE UNITED STATES

Jordan Weiss

Michel Guillot

Dementia is a leading cause of death and disability in the United States. The absolute number of adults affected by dementia is expected to rise in response to population aging, although recent studies have reported declines in the age-specific risk of dementia across sequential birth cohorts. The underlying mechanisms driving this decline remain unclear, nor is it clear whether and to what extent these trends will sustain in the future. Understanding these trends in the context of demographic and epidemiological forces would provide valuable information about the current landscape of dementia and its future burden. Microsimulation is a modeling technique that lends itself to this objective as it allows for studying life course dynamics in a counterfactual framework. In addition, it can be used to generate population forecasts in a stochastic rather than deterministic way, overcoming the untenable assumptions on which many extrapolation-based methods rely. This dissertation has three parts. In the first part, I use the parametric g-formula to examine the risk of dementia under hypothetical interventions on its risk factors while accounting for the competing risk of death. I find that observed rises in educational attainment and reduced cardiovascular disease reduce the cumulative risk of dementia but this reduction is offset after accounting for expected increases in diabetes and obesity. The second part develops a probabilistic approach to forecasting dementia incidence, prevalence, and mortality through the year 2050 while accounting for compositional and

epidemiological changes in the population. I estimate a more than two-fold increase in the number of adults aged 65 years or over with dementia from 6.3 million in 2010 to 16.3 million in 2050. In the final part, I forecast the economic burden of dementia by combining estimates from part two with linked Medicare claims and self-reported out-of-pocket expenditures. I estimate that the direct, incremental out-of-pocket expenditures and costs to Medicare for dementia may rise from $145.12 billion in 2010 to $378.98 billion by 2050.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Dementia is a leading cause of death and disability in the United States, affecting more than 5.8 million adults in 2018 with associated healthcare costs approaching $290 billion per year (Alzheimer's Association 2019). Vital statistics rank dementia as the third most common cause of death in the US (Kramarow and Tejada-Vera 2019) and it remains the only cause of death in the top 10 without a viable prevention or cure. The burden of dementia is projected to rise in response to population aging (Hebert et al. 2013) and will be especially pronounced among older adults from diverse racial/ethnic groups who are at increased risk for dementia (Babulal et al. 2019a; Haan et al. 2003; Mayeda et al. 2016; Rajan et al. 2019) and will make up an increasing share of the US population over the age of 65 in the coming decades (Colby and Ortman 2017).

The overarching objective of this dissertation is to explore how demographic and epidemiological forces shape the current and future burden of dementia.

In Chapter 1, I use a counterfactual standardization technique to examine the risk of dementia under hypothetical interventions on its risk factors. This technique, referred to as the parametric g-formula, allows users to draw population-level inference from individual-level effect estimates, making it suitable for use with microdata and appealing to demographers. Recent trends suggest a decline in the age-specific incidence and prevalence of dementia, yet the set of risk factors explaining this downward trend remain unidentified. Further, it remains unclear whether and to what extent this trend will continue. Using a potential outcomes framework, I examine the implications of shifting the population distribution of educational attainment and comorbid conditions for

dementia at the population-level while accounting for time-varying confounding and the competing risk of death.

Whereas Chapter 1 focuses on recent trends, in Chapters 2 and 3 I examine how accounting for these risk factor trends may affect the future size of the population with dementia and its associated costs, respectively.

Inherent to most studies projecting the size of the population with dementia is the assumption that age-, sex-, and race/ethnicity-specific prevalence rates of dementia will remain constant over time. This assumption attributes projected increases of the scale of the dementia burden to changes in the age structure of the population alone and ignores important risk factor trends observed among current cohorts as well as those expected among future cohorts. Dementia prevalence rates are unlikely to remain constant over time and, in fact, have reportedly declined in recent decades (Langa et al. 2017). The prevalence of dementia is affected by both the rate at which new cases occur (i.e., incidence) and the average life expectancy with dementia. Thus, changes in the underlying risk factor distribution for dementia could yield changes in its incidence. Changes in survival with dementia, independent of or together with changes in incidence, could also produce changes in prevalence rates. Accounting for these observed time trends could improve the accuracy of existing projections and aid in policy and public health planning efforts.

In Chapter 2, I develop a microsimulation model with a dementia progression module to forecast the size of the population with dementia through the year 2050 while accounting for compositional and epidemiological changes in the population. Importantly, I use survey data with Medicare linkage which allows for the inclusion of

patients with undiagnosed and diagnosed dementia. This reduces the likelihood of misclassification or undercounting of dementia cases in the population which may be especially problematic among adults whose cognitive impairment is not yet recognized by their healthcare system, perhaps a reflection of their access to healthcare, how they utilize services, and their interactions with healthcare systems.

The high economic cost associated with dementia is a fiscal burden that will increase as the size of the population with dementia rises. Understanding the magnitude of these costs and how they vary across levels of cognitive impairment is important for policymakers and healthcare systems; yet, current estimates of the incremental direct medical care costs attributable to cognitive impairment vary widely.

In Chapter 3, I provide contemporary estimates of the incremental direct medical costs of cognitive impairment without dementia (CIND), undiagnosed dementia, and diagnosed dementia. Specifically, I estimate out-of-pocket expenditures obtained through survey data and costs borne by Medicare among adults enrolled in fee-for-service (FFS) obtained through administrative data linkage. I use an econometric approach introduced by Basu and Manning (Basu and Manning 2010) to estimate costs in the presence of censoring (e.g., mortality, exit from Medicare FFS). The Basu and Manning estimator accounts for differential survival between individuals with and without cognitive impairment and can be used to estimate how costs vary from onset to death. This approach can also be used to separate the total marginal cost of cognitive impairment into an intensity of treatment and a survival component. The intensity component, which reflects the marginal cost of cognitive impairment among survivors, corresponds to variation in the accumulation of costs over time among patients with a given level of

cognitive impairment compared to those who do not have cognitive impairment. The survival component represents the incremental cost of cognitive impairment associated with cognitive impairment-specific mortality. Together, the intensity and survival components shape the total expenditure attributable to cognitive impairment by severity.

I then combine cost estimates from Chapter 3 with forecasted population counts from Chapter 2 to estimate the future direct costs of dementia in the US. To do so, I multiply the subgroup- and time-specific costs per individual with the number of individuals in each of the three impairment categories (i.e., CIND, undiagnosed dementia, dementia). In doing so, I provide new estimates of the direct economic burden of cognitive impairment through the year 2050 that account for compositional (e.g., proportion of racial/ethnic minorities, distribution of educational attainment) and epidemiological (e.g., risk of diabetes) changes in the population.

CHAPTER 1: Cumulative risk of dementia under hypothetical interventions on its risk factors: An application of the parametric g-formula

1.1 Introduction

Demographic projections of the United States indicate that the proportion of adults aged 65 years and older will increase from 13.9% in 2012 to 22.1% by 2050 (Ortman, Velkoff and Hogan 2014). This shift in the age structure of the population has major implications for age-related conditions such as dementia, a multi-etiological syndrome characterized by cognitive deficits that impair daily functioning and independence. Dementia is a leading cause of death and disability in the US (Alzheimer's Association 2019) and its risk approximately doubles every five years after the age of 65 (Jorm and Jolley 1998). The prevalence of Alzheimer's dementia, the most common form of dementia, alone is expected to increase more than 2-fold from its level in 2010 to affect an estimated 13.8 million adults by 2050, with costs in excess of $1.1 trillion (in 2018 dollars) per year (Alzheimer's Association 2019). However, in recent years, these projections—which are strictly based on changes in the age structure of the population—have been challenged by moderately consistent evidence suggesting declines in the age-specific risk of dementia.

Several observational studies in the US and Europe have reported declines in the incidence and prevalence of dementia across sequential birth cohorts over the past several decades (Derby et al. 2017; Ganguli 2017; Langa et al. 2017; Larson and Langa 2017; Prince et al. 2016; Roehr et al. 2018; Satizabal et al. 2016; Skoog 2016; Wu et al. 2017a; Wu et al. 2017b). For example, Langa and colleagues (Langa et al. 2017) observed a decline in the age-standardized prevalence of dementia from 11.6% in 2000 to 8.8% in 2012 using data from the nationally representative, US-based Health and Retirement

Study. Satizabal and colleagues (Satizabal et al. 2016) examined dementia incidence over three decades spanning the late 1970s through the early 2010s among persons aged 60 years or older in the US-based Framingham Heart Study and reported a downward trend across each successive decade. Others, including Derby (Derby et al. 2017), also reported a decline in dementia incidence in a birth cohort analysis of persons aged 70 years or older enrolled in the Einstein Aging Study from 1993 through 2015. Collectively, these studies suggest that the underlying risk of dementia may have declined in recent decades; however, the factors contributing to this decline have not been fully identified.

To date, the set of risk factors explaining the observed downward trend in dementia risk remain unidentified (Wu et al. 2017a). Instead, it is widely hypothesized that the reduced risk of dementia in later life stems from societal changes and medical advances that have improved health across the life course. This includes, for example, secular increases in educational attainment and improved treatment of cardiovascular symptomology—and particularly declines in stroke incidence—which have been consistently linked with reduced risks of dementia (Alley, Suthers and Crimmins 2007; Crimmins et al. 2018; Kalaria 2012; Livingston et al. 2017; Ryan and Bauman 2016; Sharp and Gatz 2011; Wilson et al. 2009; Xu et al. 2016). Although these trends are promising, the underlying causes warrant further study and research has yet to determine the extent to which the downward trend will sustain (Jones and Greene 2016). The marginal return of increasing educational attainment at the population level may plateau (Freedman et al. 2013), and increasing rates of diabetes and obesity (Leon and Maddox 2015) which are independently and jointly associated with increased risks of dementia (Gudala et al. 2013; Livingston et al. 2017; Whitmer et al. 2008) could plausibly stabilize

or reverse the observed decline in dementia incidence. Thus, to understand the future burden of dementia in the absence of a disease-modifying strategy, it is important to consider the underlying risk factors that could alter its trends.

Numerous studies have established individual-level risk and protective factors for dementia within a life course framework (Livingston et al. 2017). This body of work tends to focus on quantifying the associations between risk factors and dementia, which may provide insight into how they are related but does not directly address how population-level changes in the distribution of these risk factors could influence population-level trends in the risk of dementia. Establishing how, for example, increases in educational attainment at the population level co-occurring with rising levels of diabetes may relate to changes in the incidence of dementia may be more useful for informing public health initiatives than understanding how a per-unit change in a given risk factor would increase an individual's risk of dementia. Few studies (Barnes and Yaffe 2011b; Hurd, Martorell and Langa 2015; Norton et al. 2014) have examined the population-level effects on dementia of modifying its risk factors and those that have done so generally use an analytic framework that assumes the complete elimination of the underlying risk factor (e.g., population attributable risk) and do not account for the implications it may have for competing events, such as mortality. In addition, these studies tend to use traditional regression techniques which cannot account for time-varying mediators and confounders while appropriately adjusting for dynamic temporal relationships between risk factors within a counterfactual framework; an approach that has the demonstrated the advantage of mechanism investigation (Young et al. 2011).

Advances in statistical science and computational efficiency in recent decades have expanded the suite of methods used in causal inference to address these complexities and characterize the burden of disease. This includes the parametric g-formula, a generalized form of standardization that provides unbiased estimates in the presence of time-varying covariates under the assumptions of no unmeasured confounding, no measurement error, and no model misspecification (Robins 1986; Robins et al. 2009; Young et al. 2011). The parametric g-formula, like other causal inference approaches (Greenland, Pearl and Robins 1999), relies on a counterfactual framework in which the researcher can estimate potential outcomes under hypothetical treatment regimes, such as the risk of coronary heart disease under hypothetical interventions on lifestyle factors (e.g., smoking, physical activity) in the presence of competing events (Taubman et al. 2009). Thus, it allows researchers to answer questions such as how the risk of a disease would have changed in a study population had a particular risk factor been intervened upon and set to a different level than what was observed in the presence of time-varying covariates (Robins, Hernán and Siebert 2004). This method has been used in prior observational studies to assess and compare hypothetical dynamic treatment regimes in the context of heart (Taubman et al. 2009) and respiratory disease (Garcia-Aymerich et al. 2013), stroke (Vangen-Lønne et al. 2018), and diabetes (Danaei et al. 2013).

The aim of this study was to estimate the risk of dementia under hypothetical interventions on its risk factors while accounting for changes in the counterfactual risk of death. Intervention scenarios that exemplified leading theories on drivers of reductions in the cumulative risk of dementia were implemented. I applied the parametric g-formula to

observational data from the nationally representative and longitudinal Health and Retirement Study and estimated the cause-specific cumulative incidence (i.e., subdistribution function) of dementia in the presence of the competing risk of death. Compared with conventional techniques, the parametric g-formula can be used to appropriately adjust for time-dependent mediators and confounders affected by prior components of an exposure (e.g., education) in the presence of competing events (Hernán, Hernández-Díaz and Robins 2004; Naimi and Tchetgen Tchetgen 2015; Robins 1986). Here, the parametric g-formula is used to obtain estimates of the answer to the question: "How would shifting the population distribution of risk factors affect the cumulative risk of dementia in the population?"

1.2 Materials and Methods

1.2.1 Data

Data come from nine waves (2000–2016) of the Health and Retirement Study (HRS) which has been described in detail elsewhere (Sonnega et al. 2014). Briefly, the HRS is a nationally representative longitudinal study of US adults over the age of 50 with survey assessments every two years since 1992. The HRS collects a wide range of sociodemographic, health, and financial information on its participants with response rates greater than 85% at every survey wave (Sonnega et al. 2014). Respondent-level community-based sampling weights were calculated and provided by the HRS to adjust for the complex sampling design and nonresponse, which allows estimates to be generalized to the US population over the age of 50 (Heeringa and Connor 1995). The HRS is sponsored by the National Institute on Aging (NIA; U01AG009740) and is conducted by the University of Michigan.

This study was restricted to community-dwelling self-respondents over the age of 50 and dementia-free with complete data on covariates when interviewed in 2000, which is the earliest year consistent cognitive information became available. Missing values of covariates in later survey waves were carried forward from the previous survey wave. Detailed counts of exclusion criteria are shown in Table A.1.

1.2.2 Measures

1.2.2.1 Incident Dementia

Cognitive status was assessed at each wave using tests adapted from the Telephone Interview for Cognitive Status (TICS) (Ofstedal, Fisher and Herzog 2005). Total scores of cognitive status ranged from 0-27 and were derived from tests of immediate word recall (0–10 points), delayed word recall (0–10 points), serial 7s (0–5 points), and backwards counting from 20 (0–2 points). Higher scores reflect better cognitive functioning. The Langa-Weir algorithm (Crimmins et al. 2011) was applied to these continuous scores to classify participants with (scores of 7 to 27) or without dementia (scores of 0 to 6) using validated cutpoints from the Aging, Demographics, and Memory Study (ADAMS) (Crimmins et al. 2011; Langa et al. 2005). The ADAMS is a substudy of the HRS that involves 3-4 hour in-home neuropsychological and clinical assessments as well as expert clinician adjudication to obtain a gold-standard diagnosis of dementia (Langa et al. 2005).

1.2.2.2 Mortality Ascertainment

Mortality information was ascertained via survey linkage to the National Death Index through 2011 and exit interviews with surviving family members of HRS decedents thereafter with a combined 99% coverage (Weir 2016).

1.2.2.3 Educational Attainment

The primary (time-fixed) exposure in this study was participant's self-reported educational attainment (less than high school or GED, high school graduate, completed some college, college graduate, completed more than a college degree).

1.2.2.4 Covariates, Mediators, and Confounders

Covariates were selected for their established associations with dementia and availability in the HRS. Demographic characteristics included the respondent's sex (male, female), race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, or Hispanic), and maternal and paternal education (< 8 years, 8+ years). Mediators and mediator-outcome confounders included participant's self-reported height and weight which were used to calculate body mass index (BMI), level of physical activity (one or more days of vigorous activity per week or otherwise), smoking status (never smoked, former smoker, active smoker), and wealth (negative or zero; bottom tertile; middle; top tertile). Federal guidelines for BMI were used to classify respondents as underweight (BMI < 18.5), healthy weight (18.5 ≤ BMI < 25), overweight (25 ≤ BMI < 30), or obese (BMI ≥ 30) (Pi-Sunyer et al. 1998). In addition, participants self-reported at each wave whether they had diabetes, hypertension, stroke, or any heart condition (including heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems). Self-reported disease diagnoses were ascertained by asking participants to report whether a medical practitioner had ever informed them of the condition (e.g., *Has a doctor ever told you that you have diabetes or high blood sugar?*).

1.2.3 Statistical Analysis

The parametric g-formula is a generalized form of standardization that can be used to estimate an outcome distribution in the presence of time-varying covariates (Robins 1986; Robins et al. 2009; Young et al. 2011). In the current study, the parametric g-formula was used to estimate the risk of dementia under hypothetical interventions on its risk factors as a weighted sum of the probability of incident dementia conditional on past covariate histories. A formal description of the parametric g-formula is described in greater detail in the Appendix, along with a directed acyclic graph (DAG) (Greenland et al. 1999; Pearl 1995; Spirtes et al. 2000) depicting the conceptual framework of dementia—based on the literature (Alzheimer's Association 2019; Livingston et al. 2017; Patterson 2018; Wu et al. 2017a) and availability of data in the HRS—used in this study (Figure 1.1). Conceptually, the parametric g-formula is a weighted average of risks conditional on a specified intervention and observed confounder history, where the weights correspond to the probability density functions of the time-varying confounders.

To implement the parametric g-formula, I fit parametric models to the analytic HRS sample for dementia, mortality, educational attainment, and all time-varying covariates. These models were conditional on demographic characteristics, education, and lagged covariate histories in accordance with the time-varying DAG in Figure A.1.1. Individuals were considered to be at risk of incident dementia until they died or were censored at the end of follow-up in 2016; thus, loss to follow-up was not modelled. Pooled logistic models were fit for dementia and mortality and binary covariates; a sample (n=13,841) with replacement such that the empirical distribution of all baseline covariates was identical between the HRS sample and Monte Carlo generated

peudosample (step 2). Dementia and mortality were then modeled as functions of these simulated values under the natural course of exposure (i.e., observed covariate distributions) to recover the observed data distribution (i.e., the original data) and ensure adequate model specification. In step 3, the risk factor distributions were modified in accordance with specific intervention scenarios described below.

The risk of dementia—defined at age *a* as the probability an individual has developed dementia by the time they reach age *a* (Cole et al. 2015; Cole et al. 2014; Fine and Gray 1999), with age as the underlying timescale—was estimated from age 50 using subdistribution functions with a Weibull distribution. The subdistribution incidence function is an extension of the Kaplan–Meier estimator that accounts for competing risks (Lau, Cole and Gange 2009; Prentice et al. 1978; Taubman et al. 2009); in this case, mortality. Further, under hypothetical interventions, the subdistribution estimator (Gooley et al. 1999; Lau et al. 2009) accounts for changes in the counterfactual competing risk of death. That is, the model accounts for the updated risk of death under each intervention scenario. Cause-specific cumulative incidence functions are the single components that comprise a composite outcome that represents the complement of survival—in this case, the absence of dementia and all-cause mortality. The cumulative incidence function for dementia at age *a* in the absence of the competing risk of death can be represented as the complement of the survival function, $CIF(a) = 1 - S(a) = P(A_i \leq a)$, where A_i is a random variable denoting time between baseline age and incident dementia for respondent *i*. In the presence of the competing risk of death, the cause-specific cumulative incidence function for dementia at age *a* is the joint probability of either event by age *a* and that the event is dementia, $CIF_j(a) = P(A_i \leq a, J_i = 1)$,

where J_i is a random variable denoting dementia (j=1) or all-cause mortality (j=2) (Cole et al. 2014). This approach yields estimates of the population-average counterfactual risk of dementia (i.e., the proportion of individuals who develop dementia in a synthetic population with no loss to follow-up and in which individuals may die prior to incident dementia) in the presence of time-varying confounding and mediation while accounting for changes in the cause-specific cumulative incidence function for mortality that result from the hypothetical interventions.

Ninety-five percent confidence intervals were obtained by repeating steps two and three in 500 bootstrapped samples (Efron and Tibshirani 1986). Population risk ratios and risk differences were then calculated directly from the cumulative risk estimates under each intervention scenario relative to the natural course. All analyses used age as the underlying time-scale. Community-based sample weighted provided by the HRS were used to account for the complex sampling design and obtain nationally representative estimates as recommended by the HRS (Fisher and Ryan 2017; Ofstedal et al. 2016).

1.2.3.1 Intervention Scenarios

I used the parametric g-formula to estimate what the cumulative risk of dementia would have been under the following intervention scenarios implemented at the population-level. Interventions were implemented at baseline in year 2000 and continued in every wave through the end of follow-up.

(1) Increase average educational attainment level to a high school degree (Prescott 2019)

(2) Reduce risk of hypertension by 22% (Zanchetti, Thomopoulos and Parati 2015)

(3) Reduce risks of stroke by 20% (Kernan et al. 2014) and heart disease by 30% (Benjamin et al. 2017)

(4) Increase the risk of diabetes by 9.4% (Geiss et al. 2014)

(5) Increase the risk of obesity by 30% (Flegal et al. 2012)

(6) Scenarios 1-3 combined

(7) Scenarios 1-5 combined

For example, in Scenario 1, simulants were drawn with replacement until the population mean level of educational attainment was a high-school degree. Scenarios that reduced the risk of a particular condition by a certain percentage applied these reductions to all simulants regardless of their covariate distribution. In these scenarios, a proportional reduction was applied at each survey wave.

The scenarios above were chosen to exemplify population-level trends in the US based on data from the National Center for Health Statistics, National Center for Education Statistics, and from meta-analyses of randomized trials to understand how the future burden of dementia may be shaped by changes in its underlying risk factors.

Intervention scenario 6 combines rising levels of education with reductions in hypertension, stroke, and heart disease to better understand how these risk factor trends may explain the reported declines in dementia risk. Intervention scenario 7 combines higher levels of educational attainment and declines in hypertension, stroke, and heart disease with increasing levels of obesity and diabetes. This scenario may inform future trends in dementia, and whether recently observed declines, possibly due to reductions in

some risk factors (i.e., hypertension, stroke, heart disease), will be offset by increases in others (i.e., diabetes, obesity).

1.2.3.2 Sensitivity Analysis

The primary analysis estimates the hypothetical risk of dementia under numerous intervention scenarios in the presence of mortality which, too, is affected under the intervention scenarios. Intuitively, improving the educational distribution and reducing the risks of diabetes and cardiovascular disease should result in reduced mortality and, ultimately, increased risk of dementia in response to improved survival. To estimate the degree of this offset, I replicated the primary analysis for the natural course and intervention scenario 7 in the absence of mortality.

1.3 Results

Characteristics of the study sample at baseline, mid-study, and end-of-study are shown in Table 1.1. Among the 13,841 eligible participants, 54.3% were aged 50–64 years, 28.3% were aged 65–74 years, 14.5% were aged 75–84 years, and 2.9% were aged 85 years or older at baseline. At baseline, nearly 60% of participants were female; 79% were non-Hispanic white; and 27.5% had less than a high school education. Nearly 12% of these participants had a history of diabetes; 41.1% had a history of hypertension; and 24.3% were classified as obese. Over 16 years of follow-up, 1,788 (12.9%) participants developed dementia and there were a total of 6,690 (48.3%) deaths.

The covariate distributions simulated under the natural course showed close correspondence with the observed data as did the incidence functions for dementia and mortality. For example, the population-averaged difference between the observed and

simulated risk of dementia under the natural course was 0.07% (95% CI: -0.43, 0.56) over the study period.

Table 1.2 shows risk estimates (and 95% CIs) for dementia based on the parametric g-formula under the natural course and select intervention scenarios. These estimates correspond to the risk of developing dementia conditional on survival to age 50 (i.e., the proportion of dementia-free individuals at baseline who developed dementia between ages 50-105). The simulated risk of dementia under the natural course was 8.67% (95% CI: 8.52, 8.82) and similar to the observed risk at 8.74% (95% CI: 8.24, 9.23) (Figure 1.2). Population risk ratios (PRRs) and risk differences (PRDs) for each intervention scenario compared with no intervention are also presented in Table 1.2. A PRR of 1.0 indicates the intervention scenario had no effect on the risk of dementia; a PRR less than 1.0 indicates the intervention scenario reduced the risk by (1-PRR)%; and a PRR greater than 1.0 indicates that the intervention scenario increased the risk by (PRR-1)%. The PRD represents the percentage point difference between the risk of dementia under the natural course and intervention scenarios (i.e., $\text{risk}_{\text{natural course}} - \text{risk}_{\text{intervention scenario}}$).

Under a joint intervention on education, hypertension, stroke, and heart disease (Scenario 6), the risk of dementia would be 7.63% (95% CI: 6.65, 8.61); an absolute risk difference of -1.04 percentage points (95% CI: -2.02, -0.06) compared with no intervention. Combining the “beneficial” (Scenarios 1-3) and “deleterious” (Scenarios 4-5) interventions (Scenario 7) would result in a risk of dementia of 8.54% (95% CI: 6.78, 10.3) and an absolute risk difference of -0.13 percentage points (95% CI: -1.89, 1.63) compared with no intervention, eradicating much of the benefit that was observed under intervention Scenario 6. These results are illustrated in Figure 1.2 which shows the

incidence functions for the risk of dementia under the joint intervention scenarios, natural course, and using the observed data. The overlap between the natural course and observed data highlight their close fit while the shift of the incidence function under Scenario 7 towards the natural course indicates the reductions in the year risk of dementia achieved under intervention Scenario 6 were partially offset.

The counterfactual risk of dementia under the intervention scenarios is estimated in the presence of modified mortality risks. Thus, the cumulative incidence functions reflect changes in the risk of dementia that are attributable to both changes in its underlying risk factors as well as changes in the competing risk of death. This may offset some of the gains that could be seen through risk factor reduction, by improving survival outcomes and therefore extending the mean age at death in the population. To estimate the "pure" counterfactual risk of dementia under the hypothetical intervention scenarios, I replicated the primarily analysis in the absence of mortality by settings its probability to zero (i.e., effectively intervening to prevent death). Results from this analysis are presented in Figure A.1.2 and Table A.1.2. As expected, in the absence of mortality, the risk of dementia increased to 9.96% (95% CI: 9.58, 10.35) under the natural course. In this setting, the risk of dementia under Scenario 6 was 13.20% (95% CI: 12.53, 13.87); an absolute risk difference of 3.24 percentage points (95% CI: 2.96, 3.62). Under Scenario 7, the absolute risk difference was 6.89 percentage points (95% CI: 5.68, 8.19). Thus, in the absence of mortality, the difference between Scenario 6 and Scenario 7 is 3.65 percentage points compared to just 0.91 percentage points when accounting for the competing risk of mortality. Thus, assuming constant mortality, Scenario 6 would provide a much larger reduction in the overall risk of dementia.

1.4 Discussion

Despite a recent focus on the decline in the incidence of dementia, few studies have examined the underlying causes within a counterfactual framework to assert causal relations or as a means of hypothesis validation regarding its trends. Identifying causal relationships is at the core of improving our mechanistic knowledge of disease processes and informing public health initiatives and policies. In the context of dementia, a public health priority for which there is no disease-modifying strategy, understanding its future burden relies, in part, on understanding the risk factors that could alter its trends.

This study extends the body of work that examines trends in dementia by estimating the cumulative risk of dementia under hypothetical interventions on its risk factors. I used the parametric g-formula (a generalized form of standardization) to appropriately adjust for dynamic feedback loops between risk factors which lends to a more realistic modeling of the mechanistic processes affecting the cumulative risk of dementia. The simulated risk of dementia under the natural course was 8.67% (95% CI: 8.52, 8.82). This estimate is comparable to one obtained from a study in the UK-equivalent of the HRS, the English Longitudinal Study, which took into account the competing risk of mortality and structured their models with age as the underlying time scale (Rogers, Steptoe and Cadar 2017). Under a joint intervention (Scenario 6) of increased educational attainment (high school degree at the population level) and reduced hypertension (22% reduction in risk), stroke (20% reduction in risk), and heart disease (30% reduction in risk), the cumulative risk of dementia would have been reduced to 7.63 (95% CI: 6.65, 8.61); a relative reduction of 12% compared with no intervention. However, much of this reduction was eradicated when simultaneously accounting for

increasing levels of diabetes (9.4% increase in risk) and obesity (30% increase in risk) (Scenario 7) which resulted in a cumulative risk of 8.54 (95% CI: 6.78, 10.3); a relative reduction of 2% compared with no intervention.

Overall, these results link population-level improvements in risk profiles to reductions in the cumulative risk of dementia. Such results are promising and may serve as early evidence to confirm leading hypotheses that attribute the downward trend in dementia to increasing levels of educational attainment and improved cardiovascular risk factor profiles. However, the consequences of increasing levels of obesity and diabetes could be indicative of a rebound in the risk of dementia in the coming decades as reflected by the results obtained under Scenario 7. Thus, the scale of the impending dementia epidemic is likely to be influenced by efforts to prevent and control diabetes and obesity as well as initiatives that foster continued increases in educational attainment and improved treatment and prevention of cardiovascular disease.

Educational attainment in the population is expected to rise in the coming decades but there have been long standing concerns regarding if and when these gains will stagnate (Day and Bauman 2000). The incremental rises in future educational attainment are likely to occur at diminishing rates and it remains unclear if there will be a point at which population-level gains in educational attainment will be offset by increasing levels of comorbidities which appear to have a higher ceiling.

The degree of offset in dementia risk reduction observed with Scenario 7 underscores the importance of prevention and management strategies for diabetes and obesity across all ages. Lifestyle factors for diabetes (Fletcher, Gulanick and Lamendola 2002) and obesity (Hruby et al. 2016) are well-established and both feature sub-clinical

stages. Thus, despite the potential costs and enormous challenge, increased efforts for screening and early detection of at risk individuals across all stages of the lifespan may complement existing health promotion efforts (Force 2010; Gilmer and O'Connor 2010). Improved monitoring of diabetes and obesity will play an essential role in the future burden of dementia.

Dementia has been linked with a range of demographic, lifestyle, and medical risk factors in cross-sectional and prospective studies (Alzheimer's Association 2019; Livingston et al. 2017). However, many of these studies examine the association between incident dementia and a risk factor measured at baseline, or include repeated values of a risk factor in a longitudinal setting without appropriate adjustment for time-varying confounding and longitudinal dependencies. These limitations are overcome in this analysis by use of the parametric g-formula which also provides estimates of the population impact of changes in risk factor distributions rather than associations based on an existing risk exposure level.

The present study has several strengths, including its longitudinal design with up to 16 years of follow up, use of a large nationally representative sample, ascertainment of dementia and mortality status using validated criteria, and the analytical approach. The parametric g-formula allows for studying the cumulative risk of dementia under hypothetical shifts in the population distribution of its risk factors while appropriately adjusting for time-varying confounders and longitudinal dependencies, providing population-level effect estimates that are directly relevant for public health and planning purposes.

As with any observational study, the results from this analysis relied on a set of assumptions to produce unbiased estimates (Robins 1986; Young et al. 2011). Given the number of models estimated within the g-formula framework, model misspecification is a potential concern—perhaps more so than standard regression analyses, which also require correct model specification (Berry 1993), due to the number of models required. I used a DAG to specify the hypothesized pathways through which time-fixed and time-varying characteristics may affect the cumulative risk of dementia, including all relevant variables available in the HRS. In addition, I used flexible modeling techniques, such as splines, to approximate the true relationships between all variables. Adjustment for additional characteristics such as diet, occupational complexity, or other measures not available in the HRS may be necessary to account for all mediational and confounding pathways. However, the strong concordance between the observed data and simulated cumulative risk of dementia under the natural course shown in Figure 1.2, as well as the closely matched risk factor distributions for the observed and simulated characteristics as well as mortality, support the notion that this assumption held at least approximately true. Though it can be done with conventional regression approaches, studies rarely compare observational data with their simulated estimates as means of model validation which is a distinct advantage of this approach.

This study assumed counterfactual consistency which implies the counterfactual scenarios correspond to well-defined interventions (Cole and Frangakis 2009). Hypothetical intervention scenarios were designed to exemplify population-level trends and achieved reductions in risk factors observed in clinical trials. A five percent reduction in the cumulative risk of hypertension is well-defined, for example, but could be achieved

through different medications or lifestyle changes which could have different implications for the cumulative risk of dementia. Here, I assumed the change in a risk factor itself—not the means through which it was achieved—drove changes in the cumulative risk of dementia (i.e., different methods of risk factor modification would produce equivalent results).

This study also assumed the absence of uncontrolled confounding, an assumption known as conditional exchangeability (Greenland 2003; Greenland et al. 1999). The absence of unmeasured confounding implies that models account for characteristics of dementia and its risk factors; thus, any association between a risk factor and dementia is a result of the causal effect of the risk factor on dementia, rather than a common cause of the risk factor and dementia that has been excluded from the model. The assumption of conditional exchangeability cannot be tested empirically but can be mitigated to an extent by including all relevant characteristics available in a given dataset. As noted, there may be additional characteristics such as diet or occupational complexity that confound the relationships under study in this analysis. Further, it is possible that there exists a genetic variant that predisposes some individuals to a set of educational and cognitive outcomes. However, as noted by Robins (Robins et al. 2004), inference can still be drawn from the parametric g-formula in the presence of unmeasured confounding.

Positivity is another assumption of the parametric g-formula and is generally required for statistical inference (Robins et al. 2009). Positivity is met when all observed risk factor levels are observed within the joint distribution of all confounders (Westreich and Cole 2010). Positivity may arise from the use of many confounders leading to sparse data. This can pose a problem during the simulation phase as the g-formula extrapolates

over all variable combinations. However, the assumption of positivity can be evaluated empirically and was met in this study.

In addition to these assumptions, the parametric g-formula is subject to the g-null paradox (Robins and Hernán 2009) which specifies that the sharp null hypothesis of no treatment effect on the outcome will be falsely rejected in large sample sizes, regardless of the true associations. The associations between the risk factors in this analysis and dementia are well documented (Livingston et al. 2017); thus, the g-null paradox is less of a concern in this study.

Irrespective of the limitations, these findings extend the body of work documenting trends in dementia by investigating potential mechanistic pathways through which these trends may emerge. The risk of dementia is impacted by complex and interrelated pathways that are woven together over the life course. Thus, it is necessary to employ a methodological framework that allows for such dynamic processes to be modeled appropriately. Under the specified assumptions, the parametric g-formula provides unbiased effect estimates. In violation of these assumptions, the parametric g-formula still provides a representation of the dynamic and complex pathways that shape the risk of dementia. The estimates obtained in this analysis are appealing because they correspond to feasible interventions and reflect population-level impacts, providing intuitive results for public health officials and policy makers which correspond to steps that can be taken to minimize the future burden of dementia in the US.

1.5 Tables

Table 1.1 Sample characteristics at baseline (2000), mid-study (2008), and end-of-study (2016), HRS

	Survey Year		
Characteristic, n (%)	2000	2008	2016
Age			
50-64	7,521 (54.34)	2,481 (23.57)	0
65-74	3,913 (28.27)	4,767 (45.28)	2,967 (41.49)
75-84	2,003 (14.47)	2,440 (23.18)	3,127 (43.73)
85+	404 (2.92)	840 (7.98)	1,057 (14.78)
Sex			
Male	5,546 (40.1)	4,041 (38.4)	2,591 (36.2)
Female	8,295 (59.9)	6,487 (61.6)	4,560 (63.8)
Race/ethnicity			
Non-Hispanic white	10,939 (79.0)	8,354 (79.4)	5,623 (78.6)
Non-Hispanic black	1,720 (12.4)	1,266 (12.0)	875 (12.3)
Non-Hispanic other	251 (1.9)	189 (1.8)	139 (1.9)
Hispanic	931 (6.7)	719 (6.8)	514 (7.2)
Father's education			
< 8 years	5,580 (40.3)	4,030 (38.3)	2,525 (35.3)
8+ years	8,261 (59.7)	6,498 (61.7)	4,626 (64.7)
Mother's education			
< 8 years	4,535 (32.8)	3,183 (30.2)	1,985 (27.8)
8+ years	9,306 (67.2)	7,345 (69.8)	5,166 (72.2)
Education			
< HS or GED	3,812 (27.5)	2,645 (25.1)	1,585 (22.2)
HS	4,511 (32.6)	3,489 (33.1)	2,367 (33.1)
Some college	2,846 (20.6)	2,217 (21.1)	1,589 (22.2)
College and above	2,672 (19.3)	2,177 (20.7)	1,610 (22.5)
Wealth			
Negative or zero	818 (5.9)	589 (5.6)	496 (6.9)
Bottom tertile	4,343 (31.4)	3,322 (31.6)	2,226 (31.1)
Middle	4,348 (31.4)	3,312 (31.5)	2,213 (31.0)
Top tertile	4,332 (31.3)	3,305 (31.4)	2,216 (31.0)
Physically active			
Yes	7,454 (53.9)	7,427 (70.6)	4,983 (69.7)
No	6,387 (46.1)	3,101 (29.4)	2,168 (30.3)
BMI			
Underweight	200 (1.4)	173 (1.6)	167 (2.3)
Normal weight	4,835 (34.9)	3,309 (31.4)	2,281 (31.9)
Overweight	5,439 (39.3)	3,962 (37.6)	2,639 (36.9)
Obese	3,367 (24.3)	3,084 (29.3)	2,064 (28.9)
Smoking status			
Never	5,639 (40.7)	4,470 (42.5)	3,224 (45.1)
Former	5,885 (42.5)	4,861 (46.2)	3,407 (47.6)
Active	2,317 (16.7)	1,197 (11.4)	520 (7.3)
History of diabetes*	1,633 (11.8)	2,840 (20.5)	3,699 (26.7)
History of hypertension*	5,690 (41.1)	8,086 (58.4)	9,445 (68.2)
History of stroke*	672 (4.9)	1,323 (9.6)	1,951 (14.1)
History of heart disease*	2,445 (17.7)	4,038 (29.2)	5,352 (38.7)
N	13,841	10,528	7,151
Dementia cases*	0	1,095 (7.9)	1,788 (12.9)
Deaths*	0	3,313 (23.9)	6,690 (48.3)

Notes. Unweighted *n* and weighted proportions are shown. BMI, body mass index; GED, general education development test; HS, high school.

* Cumulative counts are shown in reference to the baseline sample

Table 1.2 Simulated risk and 95% CI estimates of all-cause dementia under select hypothetical intervention scenarios, HRS

Intervention	Risk of dementia	Population risk ratio	Population risk difference
(0) No intervention[*, 1]	8.67 (8.52, 8.82)	1.00	0.00
(6) Scenarios 1-3 combined[2]	7.63 (6.65, 8.61)	0.88 (0.76, 0.99)	-1.04 (-2.02, -0.06)
(7) Scenarios 1-5 combined[3]	8.54 (6.78, 10.3)	0.98 (0.78, 1.19)	-0.13 (-1.89, 1.63)

Notes. Values represent population-averaged estimates. CI, confidence interval.

[*] Observed risk 8.74 (95% CI: 8.24, 9.23).

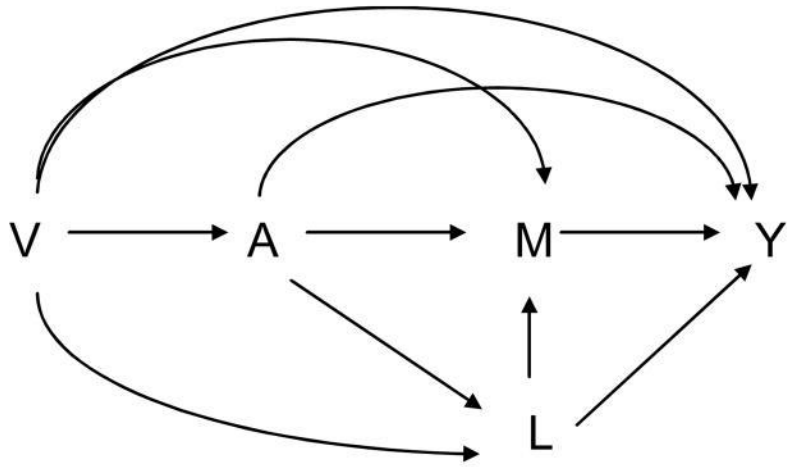
[1] Reference category.

[2] Increase average educational attainment level to a high school degree; reduce risk of hypertension by 22%; reduce risks of stroke by 20% and heart disease by 30%.

[3] Increase average educational attainment level to a high school degree; reduce risk of hypertension by 22%; reduce risks of stroke by 20% and heart disease by 30%; increase the risk of diabetes by 9.4%; Increase the risk of obesity by 30%.

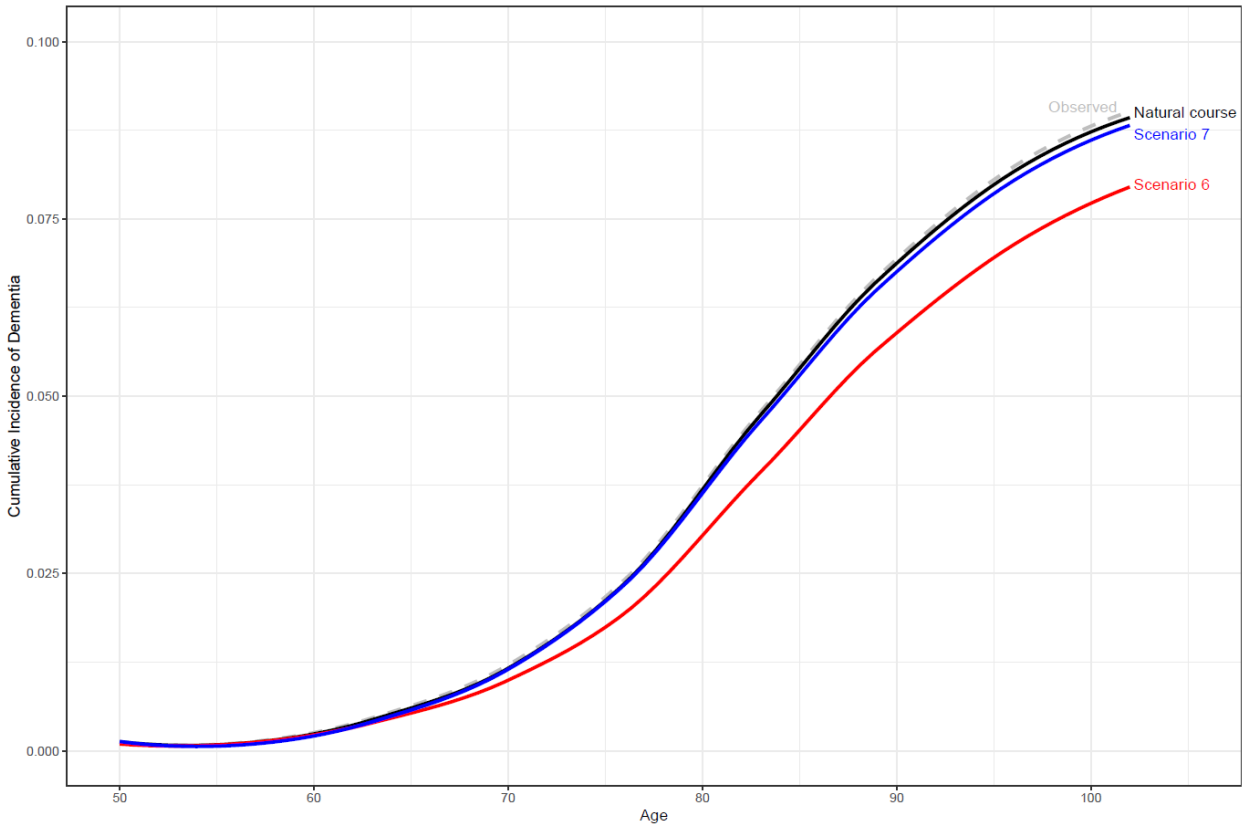
1.6 Figures

Figure 1.1 Simplified directed acyclic graph (DAG)



Notes. Represents the conceptual relationships between background characteristics V (sex, race, ethnicity, parental education), exposure A (education), mediators M (diabetes, hypertension, stroke, heart disease), mediator-outcome confounders L (physical activity, obesity, smoking status, wealth), and outcome Y (incident dementia).

Figure 1.2 Cumulative incidence curves of dementia



Notes. Observed data (dashed gray) and under the natural course (solid black) and joint intervention scenarios (Scenario 6 [solid red] and Scenario 7 [solid blue]).

1.7 Appendix

1.7.1 Technical Appendix

In the context of the current study, let k denote age with $k = 50, \ldots, j$ representing the minimum baseline and maximum attained ages among HRS respondents in the analytic sample. Let $\boldsymbol{V}$ represent a vector of baseline covariates; A represent respondent's educational attainment; $\boldsymbol{M}_k$ represent a vector of mediators at age k; and $\boldsymbol{L}_k$ represent a vector of confounders at age k. Let Y_k be an indicator of incident dementia at age k and D_k be an indicator of death at age k. Overbars indicate observed histories for time-varying variables. The risk of dementia through age j can be modeled as the probability of $Y_k = 1$ conditional on the sum of all risks incurred over all covariates, exposure, and their joint distributions.

$$\sum_{k=50}^{j} \sum_{v} \sum_{a} \sum_{\bar{m}_j} \sum_{\bar{l}_j} \left\{ \begin{array}{l} \Pr\left[Y_k = 1 \middle| V = v, A = a, \bar{M}_k = \bar{m}_k, \bar{L}_k = \bar{l}_k, \bar{Y}_{k-1} = \bar{D}_k = 0, 50 \leq k\right] \\ \times \prod_{j=50}^{k} \left[\begin{array}{l} \Pr\left[D_j = 0 \middle| V = v, A = a, \bar{M}_j = \bar{m}_j, \bar{L}_j = \bar{l}_j, \bar{Y}_{j-1} = \bar{D}_j = 0, 50 \leq j\right] \times \\ f\left[M_j = 0 \middle| V = v, A = a, \bar{M}_{j-1} = \bar{m}_{j-1}, \bar{L}_{j-1} = \bar{l}_{j-1}, \bar{Y}_{j-1} = \bar{D}_{j-1} = 0, 50 \leq j\right] \times \\ f\left[L_j = 0 \middle| V = v, A = a, \bar{M}_{j-1} = \bar{m}_{j-1}, \bar{L}_{j-1} = \bar{l}_{j-1}, \bar{Y}_{j-1} = \bar{D}_{j-1} = 0, 50 \leq j\right] \times \\ f(V = v) \times \\ f(A = a) \times \\ \Pr\left[Y_{j-1} = 0 \middle| V = v, A = a, \bar{M}_{j-1} = \bar{m}_{j-1}, \bar{L}_{j-1} = \bar{l}_{j-1}, \bar{Y}_{j-2} = \bar{D}_{j-1} = 0, 50 \leq j-1\right] \end{array} \right] \end{array} \right\}$$

This expression represents a generalized standardization procedure in the presence of time-varying covariates and exposures (i.e., the g-formula) that sums over all possible covariate and exposure histories and their joint distributions. The first probabilistic statement, indexed by k and shown at the top of the expression, reflects the time scale for risk in the current study (i.e., age). The cumulative product (indexed by j and preceded by $\Pi_{\mathrm{j}=50}^{k}$) by which that probability is multiplied represents the cumulative risk through age k for each probability of interest. This expression cannot be directly computed

nonparametrically in a high dimensional setting; thus, Monte Carlo simulation is used for the purpose of parametric approximation, whereby the parameters are drawn from the parametric equations described in Table A.1.3.

The above expression is approximated using the parametric g-formula with the actual covariate and exposure distributions which provides the natural course simulation. In the intervention scenarios, the values of $A, \bar{\boldsymbol{M}}_k, and\ \bar{\boldsymbol{L}}_k$ are "intervened on" if the predicted value obtained from the given parametric model is inconsistent with the intervention scenario by replacing the predicted value with a value that yields the desired intervention scenario in the population, on average. Then, the competing risks of dementia (Y_k) and death (D_k) are simulated probabilistically through age $k = j$ by predicting the age-specific probabilities at each age k using the above expression. This simulated dataset is then used to estimate the cumulative risk under the natural course and each intervention scenario using a cumulative incidence estimator defined below.

$$\sum_{k=50}^{j} \Pr[Y_k = 1 \mid V = v, A = a, M_k = m_k, L_k = l_k, Y_{k-1} = D_{k-1} = 0] \times \Pr[\bar{Y}_{k-1} = 0, \bar{D}_k = 0]$$

In this expression, the cumulative incidence is averaged over the simulants and summed over the covariate and exposure histories, with the frequencies of the covariate and exposure histories used as the weights. This yields the cumulative incidence function under each intervention scenario estimated via the g-formula. In a simple case where the exposure of interest is a binary indicator of attaining a high-school level of education or high (1) or otherwise (0), and the desired intervention scenario is that all individuals in the population have attained at least a high-school level of education, this process would work by (i) estimating the predicted value of educational attainment using the specified

parametric model; (ii) ensuring the predicted value is ‘1’ or, if not, replacing it with ‘1’; and then using this intervened upon value to simulate the distribution of the outcome of interest.

1.7.2 Tables

Table A.1.1 Participant exclusions, HRS

Respondents to the 2000 HRS survey wave	**19,579**
Exclusions	
Age < 50 years at baseline	554
Prevalent dementia at baseline	1,526
Lived in nursing home at baseline	107
Surveyed by proxy at baseline	1,252
Invalid sampling weight at baseline	369
Missing baseline covariates, mediators, or confounders	1,930
Excluded cases	**5,738**
Analytic sample	**13,841**

Table A.1.2 Simulated risk and 95% CI estimates of all-cause dementia under select hypothetical intervention scenarios in the absence of mortality, HRS

Intervention	Risk of dementia	Population risk ratio	Population risk difference
(0) No intervention[*, 1]	9.96 (9.58, 10.35)	1.00	0.00
(6) Scenarios 1-3 combined[2]	13.20 (12.53, 13.87)	1.33 (1.29, 1.36)	3.24 (2.96, 3.62)
(7) Scenarios 1-5 combined[3]	16.85 (15.26, 18.42)	1.70 (1.57, 1.78)	6.89 (5.68, 8.19)

Notes. Values represent population-averaged estimates. CI, confidence interval.

[*] Observed risk 10.05 (95% CI: 9.51, 10.56)

[1] Reference category.

[2] Increase average educational attainment level to a high school degree; reduce risk of hypertension by 22%; reduce risks of stroke by 20% and heart disease by 30%.

[3] Increase average educational attainment level to a high school degree; reduce risk of hypertension by 22%; reduce risks of stroke by 20% and heart disease by 30%; increase the risk of diabetes by 9.4%; Increase the risk of obesity by 30%.

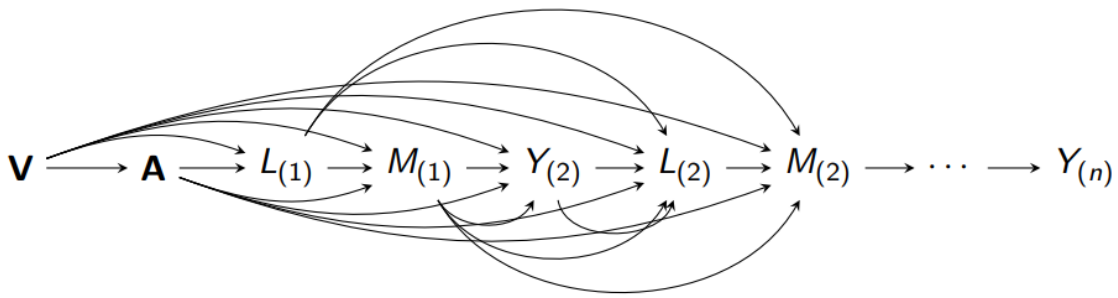
Table A.1.3 Model forms, HRS

Regrassand	Model form	Regressors
Education	Pooled logistic	sex, race, ethnicity, parental education
Diabetes	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, physical activity*, obesity*, smoking status*, wealth*
Hypertension	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, physical activity*, obesity*, smoking status*, wealth*
Stroke	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, physical activity*, obesity*, smoking status*, wealth*
Heart disease	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, physical activity*, obesity*, smoking status*, wealth*
Physical activity	Pooled logistic	sex, race, ethnicity, parental education, respondent's education
Obesity	Pooled logistic	sex, race, ethnicity, parental education, respondent's education
Smoking status	Pooled logistic	sex, race, ethnicity, parental education, respondent's education
Wealth	Pooled linear	sex, race, ethnicity, parental education, respondent's education
Dementia	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, diabetes*, hypertension*, stroke*, heart disease*, physical activity*, obesity*, smoking status*, wealth*
Mortality	Pooled logistic	sex, race, ethnicity, parental education, respondent's education, diabetes*, hypertension*, stroke*, heart disease*, physical activity*, obesity*, smoking status*, wealth*, dementia*

Notes. * Indicates a lagged term was used.

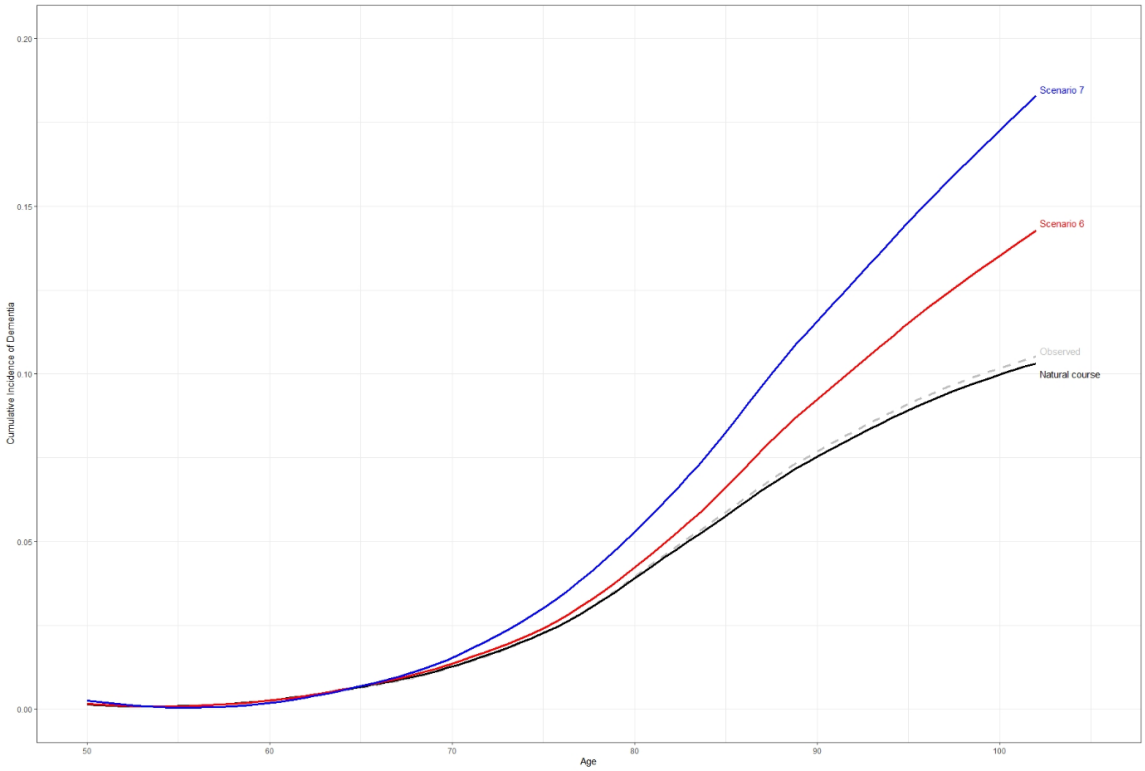
1.7.3 Figures

Figure A.1.1 Time-varying directed acyclic graph (DAG)



Notes. Represents the conceptual , time-dependent relationships between background characteristics V (sex, race, ethnicity, parental education), exposure A (education), mediators M (diabetes, hypertension, stroke, heart disease), mediator-outcome confounders L (physical activity, obesity, smoking status, wealth), and outcome Y (incident dementia). The competing risk of mortality is not depicted.

Figure A.1.2 Cumulative incidence curves of dementia in the absence of mortality



Notes. Observed data (dashed gray) and under the natural course (solid black) and joint intervention scenarios (Scenario 6 [solid red] and Scenario 7 [solid blue]).

CHAPTER 2: Dementia Population Cost Model

2.1 Introduction

Population projections have long been used to examine demographic processes, plan for the future, and to inform decision making in a variety of public health and policy contexts. In recent decades, projections focused on age-related conditions, such as Alzheimer's disease and related dementias (ADRD), have received increased attention driven by concerns about the long-term social and economic implications of population aging. Despite a growing body of literature reporting declines in the incidence and prevalence of dementia (Derby et al. 2017; Langa et al. 2008; Larson and Langa 2017; Rocca et al. 1998; Satizabal et al. 2016; Wu et al. 2017b), the absolute number of persons living with dementia is expected to rise dramatically in response to population aging and changes in population composition (Jones and Greene 2016).

To date, nearly all projections for the impending burden of dementia attribute future increases in the size of the population with dementia to changes in the age structure of the population (Brookmeyer, Gray and Kawas 1998; Hebert et al. 2003; Hebert et al. 2013; Matthews et al. 2018; Prince et al. 2016). Embedded within these estimates is the assumption that age-, sex-, and race/ethnicity-specific prevalence rates of dementia will remain constant over time which may yield inaccurate projections (Norton, Matthews and Brayne 2013). Despite the merits and utility of the existing projections, the literature reporting historical changes in both the incidence and prevalence of dementia sheds doubt on the validity of this assumption (Wu et al. 2017b). Changes in the underlying risk factor distribution for dementia could yield changes in its incidence, which could result in changes in age-specific prevalence rates of dementia. Age-specific prevalence rates could

also vary in response to changes in survival with dementia, independent of or together with changes in incidence. Thus, accounting for time trends in dementia risk factors could improve the accuracy of existing projections and aid in policy and public health planning efforts. This study introduces and validates a probabilistic microsimulation model to forecast the future burden of dementia through 2050 while accounting for demographic (e.g., population aging, racial composition) and epidemiological (e.g., cardiovascular disease, diabetes) forces in the population.

2.1.1 Dementia as a Public Health Priority

Dementia is a public health priority and its importance is expected to increase in coming decades in response to population aging. An estimated 5.6 million adults in the US age 65 and older were living with Alzheimer's disease (AD) dementia in 2019 (Alzheimer's Association 2019). AD, a degenerative brain disease characterized by progressive accumulation of the protein fragment beta-amyloid (i.e., plaques) outside neurons in the brain and twisted strands of the protein tau (i.e., tangles) inside neurons (Hardy and Higgins 1992; Selkoe 2001), is the sixth-leading cause of death (Kochanek et al. 2016)—fifth among adults age 65 and older (Doblhammer et al. 2014)—and chief cause of dementia in the US (Alzheimer's Association 2019).

Dementia is a syndrome characterized by progressive cognitive decline, such as difficulty with language or memory, that impairs an individual's social function and activities of daily living (ADLs), such as bathing, dressing, and grooming (Karlawish 2011). Consequently, persons with dementia may experience lower levels of well-being, poorer quality of life, and increased vulnerability. Dementia is not considered a part of normal aging, and its burden extends beyond the patient to their caregivers and family

while exacting a steep toll with respect to health and long-term care systems (Alzheimer's Association 2019).

There are several types of dementia, such as AD which accounts for 60% to 80% of dementia cases (Alzheimer's Association 2019), vascular dementia which is caused by vascular disease and accounts for an estimated 10% of dementia cases (Alzheimer's Association 2019), and dementia of mixed etiology where patients show signs of both AD and vascular disease (approximately 50% of AD cases) (Fernando and Ince 2004; Schneider et al. 2007). Together, AD and related dementias (ADRD) affected an estimated 14.7% of adults over the age of 70 in 2010. The individuals who help provide care for these persons make up a growing population of more than 16 million informal caregivers who, in 2018, provided an estimated 18.5 billion hours of unpaid work valued at $234 billion (Alzheimer's Association 2019). Further, caregivers face reduced quality of life (Etters et al., 2008) and increased risk of early mortality (Vitaliano et al., 2003). Thus, the widespread and significant implications dementia has on patients and their families together with larger social and economic systems—for which many countries are not fully prepared (Wortmann 2012)—underscore the importance of obtaining updated estimates of the future number of persons living with dementia.

2.1.2 Risk Factors for Dementia

It is widely believed that dementia develops as a result of multiple risk factors across numerous domains (e.g., demographic, health, lifestyle, genetic) over the life course, with potential cohort effects depending on the characteristics of the population and cultural context (Ganguli 2017; Livingston et al. 2017).

Chronological age has been implicated as the leading risk factor for dementia, with the majority of dementia cases occurring among individuals age 65 or older. A meta-analysis reported that the incidence of dementia increases exponentially through age 90, and doubles approximately every five years above the age of 65 (Jorm and Jolley 1998). Huge and Ganguli (Ganguli 2017) reported in their recent study that age was the only risk factor consistently associated with dementia at age 80 and above.

On average, women tend to be at greater risk for dementia compared to men (Podcasy and Epperson 2016) with studies indicating a divergence at around age 80 (Licher et al. 2019; van der Flier and Scheltens 2005). The extant literature on drivers of this disparity are inconclusive but several hypotheses have been put forth (Altmann et al. 2014; Babulal et al. 2019b; Lin et al. 2015; Mielke 2018; Rocca, Grossardt and Shuster 2014; Snyder et al. 2016). For example, it is believed that women's life expectancy, which is, on average, longer than men's, could contribute to the disparity in the lifetime dementia risk (Mielke 2018). A study by Chêne and colleagues (Chêne et al. 2015) observed that cardiovascular-specific mortality among men was higher than it was for women. This could serve as a selection mechanism resulting in men over the age of 65 years with comparatively better cardiovascular health than women which could protect against dementia and contribute to disparities. Studies have also reported that genetic (Altmann et al. 2014) or preclinical markers (Snyder et al. 2016) could differentially alter the risk of dementia onset in men and women.

Studies in the US have reported higher rates of dementia among non-white adults, with estimates indicating that the prevalence of dementia may be twice as high among non-Hispanic black (Gurland et al. 1999; Potter et al. 2009) and one and one-half times

higher among Hispanic adults (Gurland et al. 1999; Samper-Ternent et al. 2012) relative to age-matched non-Hispanic white adults. These disparities are believed to stem from differences in the onset risk of dementia rather than survival with dementia which remains important but of a lesser magnitude in comparison (Barnes and Bennett 2014; Chin, Negash and Hamilton 2011; Mehta et al. 2008; Wu et al. 2017b).

It is well established that lower levels of educational attainment are associated with a higher risk of dementia although the underlying mechanisms remain unclear (Sharp and Gatz 2011). The relationship between educational attainment and socioeconomic status has long been recognized and suggested as a pathway through which lower educational attainment could increase the risk of dementia (Karp et al. 2004; Yaffe et al. 2013). It has also been suggested that, together with other early-life experiences, educational attainment helps shape cognitive reserve which, it has been postulated, may account for individual differences in susceptibility to age-related brain changes or dementia-related pathology (Stern 2002, 2009, 2012).

In addition to these demographic characteristics, numerous studies have identified mid-life risk factors for dementia that include obesity and smoking. Obesity, typically determined by measuring body mass index (BMI), is hypothesized to influence dementia risk through its association with cardiometabolic disease risk (e.g., vascular disease, diabetes) (Anjum et al. 2018; Kivimäki et al. 2018). Some longitudinal studies that investigate associations between smoking and incident dementia report a protective effect (Graves et al. 1991; Hebert et al. 1992; Salib and Hillier 1997; Van Duijn and Hofman 1991). However, this is likely an artefact of the statistical methods utilized in such studies. Smoking increases one’s risk of early mortality; if one dies prior to dementia

onset, it could appear that smoking protects against dementia even after adjusting for age. Studies that examine associations between smoking and dementia while accounting for the competing risk of mortality report an increased risk of dementia among smokers (Anstey et al. 2007; Chang et al. 2012; Durazzo et al. 2014; Gorina and Kramarow 2011; Zhong et al. 2015). As with obesity and other lifestyle factors, smoking is believed to influence cognitive health and dementia through its influence on cardiometabolic disease risk (Samieri et al. 2018). While obesity and smoking are associated with increased dementia risk, physical activity has been linked to reduced risk of dementia in several meta-analyses of randomized control trials (Denkinger et al. 2012; Farina, Rusted and Tabet 2014; Groot et al. 2016).

Together with these demographic and lifestyle risk factors, medical comorbidities have also been implicated in dementia onset. It is well established that the presence of one or more cardiometabolic diseases increases dementia risk (Jefferson 2014; Kivimäki et al. 2019; Kontari and Smith 2019). This includes, for example, diabetes, stroke, and hypertension. Numerous studies have reported strong associations between a mid-life cardiovascular risk score and dementia risk in later-life (Eskelinen et al. 2009; Exalto et al. 2014; Rusanen et al. 2014; Vuorinen et al. 2015), highlighting the importance and lasting implications of mid-life cardiometabolic health for determining dementia onset at older ages.

Collectively, these risk factors and their population-level trends help shape the future burden of dementia. Thus, modeling the future burden of dementia relies on understanding and modeling underlying trends in its risk factors.

2.2 Materials and Methods

2.2.1 Data

2.2.2 Health and Retirement Study

The Health and Retirement Study (HRS) is an ongoing longitudinal panel survey of more than 37,000 US adults over the age of 50 and their spouses of any age (Sonnega et al. 2014). The HRS was designed to study changes in labor force participation and the health transitions that individuals undergo toward the end of their work lives and in the years that follow. Since its launch in 1992, the HRS has biennially collected detailed economic and health information on its participants including assets and income, mental and physical health, utilization of health services and health care expenditures, family connections, and public and private support systems.

The initial HRS cohort was recruited in 1992 using a multi-stage area probability sample designed to represent all non-institutionalized, community dwelling adults residing in households in the contiguous United States who were born in 1931-1941 (i.e., those who were between the ages of 51–61 when the study began in 1992). There was an oversampling of Blacks (1.86:1), Hispanics (1.72:1), and residents of the state of Florida (2:1) to allow for independent analysis of these subgroups. Institutionalized persons (i.e., those in prisons, jails, nursing homes, long-term or dependent care facilities) were initially excluded from the survey population which is a common practice for population surveys. However, participants were still followed if they moved from the household population into one of these institutions during the survey period.

Since 1998, the objective of the HRS has been to provide information about the US population over the age of 50 through biennial surveys. Thus, the HRS has employed

a steady-state sampling design, refreshing the original sample with a new cohort of individuals (ages 51-56 at baseline) every 6 years (i.e., 1998, 2004, 2010, 2016). The Early Baby Boomer (EBB) sample (born in 1948-1953) was added in 2004; Mid Baby Boomers (MBB; born in 1954-1959) added in 2010; and the Late Baby Boomers (LBB; born in 1960-1965) sample added in 2016. Detailed information on the sampling strategy, response rates, and overlap of these cohorts with existing HRS participants has been previously published (Sonnega et al. 2014). In addition, the HRS survey data have been supplemented with several important linkages, including administrative data from Social Security earnings and benefits records, National Death Index data, Medicare and Medicaid claims record data, and employer pension data.

2.2.2.1 Proxy Respondents

A core challenge of population-based surveys is including participants whose physical or cognitive limitations prevent them from completing an interview. Excluding these individuals on the basis of their health limitations may result in a selectively healthy sample that does not accurately reflect aging processes by underestimating the prevalence of age-associated diseases, such as dementia. The HRS researchers addressed this concern by conducting interviews with proxy informants if respondents were taking an unusually long time to complete the initial part of the survey, provided more than the threshold number of *don't know* responses, or scored below a threshold on the cognitive testing (HRS Staff 2008). Proxy interviews were also conducted in rare cases when a respondent was unavailable during the survey period, unable to complete the interview in English or Spanish, or unwilling to be interviewed but did not object to someone else answering questions on their behalf.

Proxies were primarily spouses or family members of the respondent, and were responsible for approximately 9% of interviews at each wave; proxies were used in approximately 18% of interviews for respondents age 80 and older. The relationship between the HRS participant and the proxy informant is documented in a variable in the core data files, and an additional variable is used to indicate whether the same proxy informant was the same individual for successive waves. Proxy informed interviews used a modified version of the questionnaire which involved wording changes (e.g., from "you" to "him" or "her") in most cases. Proxy informants were asked a different set of questions to assess the cognitive status of the HRS respondent. Questions to the respondent on psychological depression, cognitive testing, and subjective evaluations were omitted entirely.

2.2.2.2 Sample Weights

With its multistage national probability sampling and intentional oversampling of Blacks, Hispanics, and Florida residents, the HRS sampling design is complex. Thus, HRS investigators developed sample weights as well as clustering and stratification variables to be used for data analysis (Fisher and Ryan 2017; Heeringa and Connor 1995; Ofstedal et al. 2011). Sample weights for community dwelling respondents surveyed by self-report or by proxy are weighted together. Separate sampling weights for nursing home residents were introduced in the year 2000, making the HRS representative of non-institutionalized adults over the age of 50 and nursing home populations.

The sample weights account for the unequal probabilities of selection into the sample. Specifically, the sample weights adjust the proportion to which responses count based on the probability with which these individuals represent others in the population

as a whole. The sample weights were developed by statistically adjusting for survey nonresponse and post-stratifying the HRS sample to national population data (i.e., the US Current Population Survey through 2004 and the American Community Survey for 2006 and later waves). Not using sample weights in statistical analysis would result in oversampled individuals' responses being over-emphasized relative to non-Blacks, non-Hispanics, and non-Florida residents (Fisher and Ryan 2017).

2.2.2.3 Aging, Demographics, and Memory Study

The Aging, Demographics, and Memory Study (ADAMS), a supplement to the HRS, is an in-home neuropsychological assessment designed to provide a diagnostic determination of cognitive impairment among a subsample of HRS participants (Langa et al. 2005). ADAMS participants were drawn from a stratified, random subsample of 1,770 HRS participants ages 70 and older. Among the 1,770 eligible HRS participants, 856 individuals (48.4% of the sample) were visited by a clinical research nurse and psychometric technician, both of whom were trained in the evaluation of dementia. Interviews were conducted in the presence of an informant (e.g., family member, friend, or paid helper) and included surveys of medical history, neuropsychological testing, caregiving and health care utilization as well as their associated costs. A final diagnosis of cognitive status (normal cognitive function, cognitive impairment without dementia [CIND], and dementia) for each ADAMS respondent was established during a consensus conference by experienced teams at the Duke University Dementia Epidemiology Research Center based on guidelines from the Diagnostic and Statistical Manuals of Mental Disorders (DSM), DSM-III-R (American Psychiatric Association 1980) and DSM-IV (American Psychiatric Association 1994). Detailed descriptions of the ADAMS

sample, procedures, and assessments have been previously published (Heeringa et al. 2009; Langa et al. 2005).

2.2.2.4 Cognitive Status Among Self-Respondents

Among self-respondents, cognitive status was attained using the modified Telephone Interview for Cognitive Status (TICS) (Brandt, Spencer and Folstein 1988), a version of the Mini Mental State Examination (MMSE) (Folstein, Folstein and McHugh 1975) adapted for telephone administration. The modified TICS is an unweighted 27-point cognitive index developed from three interviewer administered items that represent short-term memory, working memory, and processing speed. The cognitive index includes: immediate and delayed 10-noun free recall tasks to measure short-term memory; a serial sevens subtraction task to measure working memory; and a backwards counting task to measure speed of mental processing. There is strong evidence to suggest these are among the first cognitive domains to decline during healthy aging (Bäckman, Small and Wahlin 2001).

2.2.2.5 Cognitive Status Among Proxy-Respondents

The short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (Jorm 1994; Ofstedal et al. 2005) was used in place of cognitive testing for HRS participants who were surveyed by proxy. Proxy respondents were asked 16 questions about the respondent's change in memory for various types of information, including the respondent's change in remembering things that have happened recently, remembering what day and month it is, and handling financial matters. All questions were worded as follows: "Compared with two years ago, how is (R's first name) at remembering things about family and friends, such as occupations, birthdays, and

addresses? Has this improved, not much changed, or gotten worse?" Each item was coded on a 5-point Likert scale from much improved (1) to much worse (5), yielding a possible score range of 16-80; higher scores indicated less improvement.

2.2.2.6 Classification of Cognitive Status

The data obtained from the ADAMS have been used to develop methods to classify HRS participants on the basis of the cognitive testing available in the larger survey (Crimmins et al. 2011; Herzog and Wallace 1997; Langa, Kabeto and Weir 2010).

Langa and Weir (Langa et al. 2010) developed cut-points to be used for the TICS that produced the same population distribution of cognitive statuses ascertained through the ADAMS. The Langa-Weir approach classifies participants who score 0-6 points as having dementia, 7-11 as CIND, and 12-27 as cognitively normal. A recent validation study found a concordance rate of 78% for dementia diagnosis when using these tests compared with the detailed ADAMS clinical evaluation (Crimmins et al. 2011).

Langa and Weir also developed an 11-point scale to assess cognitive status among HRS participants surveyed by proxy. Proxies were asked to report on the respondent's memory using a 5-point Likert scale from excellent to poor (0-5); assess whether the respondent had limitations in five instrumental activities of daily living, including managing money, taking medication, preparing hot meals, using phones, and shopping for groceries (0-5); and the survey interviewer was asked to assess whether the respondent had difficulty completing the interview because of a cognitive limitation on a scale of 3-point scale of none, some, and prevents completion (0-2). HRS participants

with scores of 6-11 were classified as having dementia, and those who scored 3-5 were classified as CIND (Crimmins et al. 2011).

The current study uses imputed measures of cognitive status provided by the HRS. Fisher and colleagues (Fisher et al. 2013) describe this imputation process in detail and note that the values are not missing at random but instead tend to be missing for those with lower levels of cognitive functioning.

2.2.2.7 Measures

2.2.2.7.1 Demographic

Demographic characteristics included the respondent's age (continuous), sex (male, female), race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, or Hispanic), and educational attainment (less than high school or GED, high school, some college, college and above).

2.2.2.7.2 Lifestyle

I included BMI and smoking status (never smoked, former smoker, active smoker). Federal guidelines for BMI were used to classify respondents as underweight (BMI < 18.5), healthy weight ($18.5 \leq$ BMI < 25), and overweight or obese (BMI ≥ 25). The decision to group overweight and obese adults was based on prior work (Fitzpatrick et al. 2009; Singh-Manoux et al. 2018; Xu et al. 2011), including a study in more than 10,000 men and women over a 28-year period (Singh-Manoux et al. 2018), which suggests adults who are overweight are more similar to those who are obese in their risk of dementia (in terms of increased risk relative to adults who are healthy weight) compared to those who are healthy weight.

2.2.2.7.3 Health

HRS participants self-reported at each wave whether they had diabetes, hypertension, stroke, or any heart condition (including heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems). Self-reported disease diagnoses were ascertained by asking participants to report whether a medical practitioner had ever informed them of the condition (e.g., *Has a doctor ever told you that you have diabetes or high blood sugar?*).

2.2.2.8 HRS-Linked Medical Claims Data

Medicare and Medicaid claims data are available for HRS participants who provided their US Social Security number. Nearly 80% of HRS participants agreed to provide this information, allowing HRS investigators to create a linked dataset comprising HRS and Medicare summary claims and utilization data (Centers for Medicare & Medicaid Services 2018). These data are not publicly available but can be attained through an application process through the HRS that involves obtaining permission from the Medicare & Medicaid Resource Information Center (MedRIC), which creates and supplies the data as well as authorization from the Research Data Assistance Center (ResDAC). The CMS records include a unique identifier for each respondent that can be linked back to the HRS data through a file supplied by the HRS investigators. At the time of this dissertation, the CMS Medicare Parts A and B Claims and Summary Files—which comprise the fee-for-service (FFS) component of the Medicare program—spanned the years 1991-2015 and were available for 26,044 participants (Centers for Medicare & Medicaid Services 2018).

A clinical diagnosis of dementia was ascertained from linked Medicare claims based on the Chronic Conditions Data Warehouse algorithm. The current study used ICD-9 codes to classify persons with dementia on the basis of their prior validation by Taylor and colleagues (Gorina and Kramarow 2011; Taylor Jr et al. 2009). The ICD-9 codes used in the current study are shown in Table A.2.1.

2.2.3 Population Data

In the US, Federal law mandates the collection and publication of vital statistics data, including information obtained from birth and death certificates which are required to be completed for all births and deaths across all states. This information is compiled and made publicly available by the National Vital Statistics System (NVSS) through a joint initiative between the National Center for Health Statistics (NCHS) and state health departments of the US to provide access to statistical information.

The US Census Bureau provides historical population estimates for each calendar year. From these data, I obtained estimated counts from the years 2000-2018 for the number of adults 51-years of age disaggregated by gender, race, and ethnicity (U.S. Census Bureau 2018).

In March 2015, the US Census Bureau released updated population projection counts through the year 2060 (Colby and Ortman 2017). These projections were produced using the cohort-component method and based on the 2010 Census and official estimates through 2013. The cohort-component method projects separately for each birth cohort the components of population change (i.e., fertility, mortality, net international migration). The base population is then updated each year using projected survival and

net international migration rates. Projected fertility rates are applied to the female population to generate successive birth cohorts which are added back in to the population and used for projections. It is important to note that these projections do not reflect current trends but instead attempt to predict future change. Whereas some projections are based primarily on historical population growth patterns, the Census Bureau prepares estimates based on population counts consistent with the most recent decennial census. These projections report the estimated number of individuals in 1-year age bands from 0-100 years of age, stratified by gender, race, and Hispanic origin. This study used the US Census 2014 National Population Projection for the period 2014 to 2060 (Colby and Ortman 2017) to obtain projections of the number of 51-year olds by gender, race, and ethnicity in each calendar year from 2019 through 2050 (U.S. Census Bureau).

2.2.4 Behavioral Risk Factor Surveillance System

The Behavioral Risk Factor Surveillance System (BRFSS) is a nationally representative telephone survey of adults in the US (Centers for Disease Control and Prevention). BRFSS is sponsored by the Centers for Disease Control and Prevention (CDC) which conducts more than 400,000 interviews each year, making it the largest continuously conducted health survey system in the world. The BRFSS is used to collect data about US adults regarding their sociodemographics, health-related risk behaviors, and chronic health conditions.

I used survey data from the BRFSS (years 2001-2018) to model and project risk profiles of replenishing cohorts as described in Section 2.2.6.5.

2.2.5 Human Mortality Database

The Human Mortality Database (HMD) is a unique, open-access collection of detailed demographic data for 40 countries with validated vital registration and census information (Barbieri et al. 2015; Human Mortality Database). The information contained within the HMD includes but is not limited to information on births, deaths, population counts, and death rates. Using data from National Statistics Offices, the HMD constructs and provides life tables by year, sex, and age using standardized methods for all time periods to ensure maximum comparability and mitigate issues related to data quality (Wilmoth et al. 2007). The current study used death rates obtained from the HMD to compare estimated death rates obtained from the simulation as described in Section 2.2.6.9.

2.2.6 Statistical Analysis

Several steps are necessary to construct the microsimulation model used to forecast the future burden of dementia. First, I estimate covariate-specific transition matrices for the state space using the HRS data from 2000 to 2010. Then, I use HRS-provided sampling weights for the year 2000 to weight the baseline sample up to the national population; this sample is used as the base population. The size and composition of replenishing cohorts are obtained from the US Census Bureau, and their corresponding risk profiles are forecasted using data from the BRFSS. The simulation process follows. In each simulation cycle, the base population and replenishing cohorts are exposed to the set of transition matrices which are used to determine if and how their movement in the state space changes. This process is continued for 50 simulation cycles spanning the years 2000 to 2050.

2.2.6.1 Base Population

An essential input to the microsimulation is the base population for the initial simulation year. The HRS is an ideal data source to use as input to the base population due to its portability, wealth of micro-level information, and because it is nationally representative of the community-dwelling and nursing home populations in the US. I used the year 2000 survey wave of the HRS to inform the base population because it marks the first HRS survey year in which cognitive status was ascertained in a consistent manner. Characteristics of the base population consist of sociodemographic, lifestyle, and health-related factors as described in Section 2.2.2.7.

The base population was restricted to adults over the age of 50 with valid sampling weights (either community-based or nursing home-based) when surveyed in the year 2000. Missing covariate values for lifestyle and health characteristics in the year 2000 survey wave were imputed using a form of multiple imputation by chained equations (Buuren and Groothuis-Oudshoorn 2010). This approach works well in practice and is suitable for imputing incomplete large, national, public datasets (Oudshoorn 1999; Van Buuren et al. 2006). Imputed values were based on age, sex, race/ethnicity, and education (there were no missing values for educational attainment). Sample weights were not used for nor prior to imputation. Unweighted sample characteristics before and after imputation are shown in Table 2.1 along with percent missing for each variable prior to imputation. The combined person-level and nursing home weights were used to parameterize the covariate distribution and composition of the base population so that it was representative of the non-institutionalized population over the age of 50 and nursing home populations. Table 2.2 presents weighted and unweighted sample characteristics of

the base population. Figure A.2.1 compares population counts between the base population (i.e., weighted HRS sample in year 2000) and the year 2000 Census by age.

Compared with the year 2000 population counts provided by the Census, the sample-weighted HRS undercounted all age groups. Thus, HRS respondents were upweighted in proportion to their HRS-provided sampling weights to match the age, sex, and race/ethnicity composition for the year 2000 derived from the Census.

2.2.6.2 State Space

The current study expanded the state-space of existing dementia projections which typically focus exclusively on AD-specific or all-cause dementia. Here, the state-space is defined as: absence of cognitive impairment (State 1); cognitive impairment without dementia (State 2); undiagnosed dementia (State 3); diagnosed dementia (State 4); and death (State 5). Undiagnosed dementia refers to individuals who are classified with dementia based on the HRS survey but for whom a diagnosis does not exist in HRS-linked Medicare claims data. Diagnosed dementia refers to individuals for whom a dementia diagnosis does exist in their linked claims data irrespective of their cognitive status based on their classification in the HRS. The state space is shown graphically in Figure 2.1.

2.2.6.3 Multistate Modeling

An individual's health trajectory can be conceptualized as a sequence of transitions between progressive states of health whereby transitions are dependent on an individual's characteristics as well as their occupied state and duration of time spent in

that state. A multistate model can be used in this context to identify and quantify the effects of risk factors associated with the different transitions between states over time.

This process can be formalized by defining the state space S as the set of finite states $\{s_1, s_2, \dots, s_k\}$ an individual can transition between. Thus, $S = \{s_1, s_2, \dots, s_k\}$ denotes the finite discrete-state space containing k states. The quantity k can be thought of as the number of outcome states for a health condition. For illustrative purposes, consider the two-state model in which $k = 2$ and $S = \{s_1, s_2\}$ represents being alive (s_1) or dead (s_2). Then, the state of the process (i.e., alive or dead) at time t can be defined as Y_t for $t = 0,1,2, \dots n$ and the discrete-time process $\{Y_t \mid t = 0,1,2, \dots, n\}$ can be written as the joint probability mass function using a product of conditional probabilities across all time points:

$$\begin{aligned}
&\Pr(Y_0 = y_0, Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n) \\
&\quad = \Pr(Y_0 = y_0) \times \Pr(Y_1 = y_1 \mid Y_0 = y_0) \times \Pr(Y_2 = y_2 \mid Y_1 = y_1, Y_0 = y_0) \\
&\quad \times \dots \times \Pr(Y_k = y_k \mid Y_{k-1} = y_{k-1}, Y_{k-2} = y_{k-2}, Y_0 = y_0).
\end{aligned}$$

If the one-step transition probability from state i to state j at time m is generalized as:

$$\Pr_{ij,m} = \Pr(Y_{m+1} = j \mid Y_m = i)$$

then the joint probability mass function $\Pr(Y_0 = y_0, Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n)$ can be defined by the one-step transition probability matrix:

$$\mathbf{Pr}_{\boldsymbol{m}} = \begin{pmatrix} \Pr_{11,m} & \dots & \Pr_{1k,m} \\ \vdots & \ddots & \vdots \\ \Pr_{k1,m} & \dots & \Pr_{kk,m} \end{pmatrix}, \text{ where:}$$

$$\sum_{j=1}^{k} \Pr_{ij,m} = 1$$

is satisfied across all rows of $\mathbf{Pr}_{\boldsymbol{m}}$. In other words, the transition from state i at time m to all states in the set $S = \{s_1, s_2, \dots, s_k\}$ at time $m + 1$ must sum to one.

These probabilities can then be written as a function of a vector of covariates $\boldsymbol{V}$, linking the transition probabilities to individual characteristics through a series of j multinomial logistic regressions defined by:

$$\log\left(\frac{\Pr_{\mathrm{ij}}}{\Pr_{ii}}\right) = \beta_{\mathrm{ij},0} + \beta_{ij}^{T}\boldsymbol{V}, \qquad i \neq j, T = \{0,1,2, \dots, t\}$$

with one model for each row of the transition probability matrix.

Typically, these quantities are estimated with a Markov chain of the first order which assumes that, given the present state s_i at time $t = m$, the transition to state s_j at

time $t = m + 1$ is independent of all past states and duration in those states (i.e., the "memoryless" property of the Markov assumption). Of course, in the context of age-related health conditions, the mean duration of occupancy in a given state (i.e., sojourn time; the average time spent in a specific state before transitioning into another state) s_i at time $t = m$ may alter the probability of transitioning into state s_j at time $t = m + 1$. For example, an 80-year old with a 10-year history of functional impairment may be more likely to transition to death than an 80-year old with a 1-year history of functional impairment. This higher order dependency implicates the importance of individual histories as defined by a sequence of previous states and time spent in those states which can be incorporated in a second- or higher-order Markovian framework. Here, I use a first-order Markovian framework.

2.2.6.4 Transition Models

Transition models were used to estimate the probability of transitioning between different states in the state space described in Section 2.2.6.2. Transition probabilities, denoted by $\Pr(Y_{m+1} = j \mid Y_m = i)$ or p_{ij}, represent the probability that an individual in state i at time m will be in state j at time $m + 1$. Markov chains are commonly used to estimate transition probabilities using longitudinal data. There are several ways in which one can estimate a Markov chain, including logistic regression. Logistic regression offers the advantage of flexibly incorporating time-varying covariates that are straightforward to interpret (Islam and Chowdhury 2006; Lawless and Rad 2015; Salazar et al. 2007; Yee 2015). Thus, I used multinomial logistic regression to estimate the probability of transitioning between states in the state space.

In the regression setting, the state j at time $m+1$ is modeled as a function of the states at times $m, m-1, m-2$, and so on to capture an individual's trajectory and time spent in each state. Age effects were modeled flexibly using splines which allow the transitions to vary non-linearly but smoothly with the aging process. In addition to age, these models included sex, race/ethnicity, education, and the complete list of lifestyle and health characteristics noted in Section 2.2.2.7.

I estimated transition probabilities across the states {1=no cognitive impairment, 2=cognitive impairment without dementia, 3=undiagnosed dementia, 4=diagnosed dementia, 5=death} by structuring the base population for the years 2000-2010 in person-period format and estimating a series of multinomial regressions. This follow-up period was selected because it contains mortality records validated through a mortality database (the National Death Index). Data for the survey years 2012 and 2014 were excluded from the estimation of transition probabilities to allow for interval validation as described in Section 2.2.6.8. I chose not to use split sampling for internal validation because it has been found to be outperformed by bootstrap resampling (Steyerberg et al. 2001). Transition times between states 1-4 were not known exactly (i.e., they were interval-censored) but the occurrence time of state 5 (i.e., death) was known.

Importantly, the HRS conducts surveys every two years. Thus, the estimated transition probabilities reflect the probability of transitioning across states in two-year increments. Annual transition probabilities can be interpolated from these two-year transition probabilities using the approximation (Chhatwal, Jayasuriya and Elbasha 2016)

$$p_{ij}^{*} = 1 - \left(1 - p_{ij}\right)^{\frac{1}{t}}$$

applied to all destination states and subtracting their sum from one to obtain the probability of remaining in the current state. In this context, $t = 2$. It is also necessary to assume only one transition occurs between each survey wave in accordance with Figure 2.1. An alternative approach to approximating annual transition probabilities would be the eigendecomposition approach which requires taking the nth root (i.e., the square-root in this context) of the two-year probability matrix. However, this approach is noted for complications arising from the diagonal elements of the transition matrix becoming negative after eigendecomposition (Chhatwal et al. 2016). Due to the nature of the modeling procedure, for some transitions, there were trivial but positive probabilities of recovering from death. These probabilities were corrected by zeroing out the probability of recovery from death and setting the probability of remaining in the death state to one.

Transition models for the state space were estimated 500 times using a bootstrap approach. Draws from the base population were made—with replacement—to account for the joint distribution of population characteristics. The transition parameters were then estimated each of these 500 times to generate uncertainty bounds for the parameter distribution. Sample transition probabilities are shown in Figures A.2.2, A.2.3, and A.2.4.

The evolution of covariate histories is observed for the base population; time-varying covariates (e.g., BMI, diabetes, hypertension) are measured every two years. However, the evolution of covariate histories is unknown for replenishing cohorts. Although it is possible to establish the distribution of risk factors in each replenishing cohort at baseline, evolving these risk factors forward in time requires an additional series of incidence models to obtain transition probabilities. Thus, in addition to the transition model for the state space, I estimate a series of models for BMI and medical risk factors;

it is assumed that smoking status at age 51 is fixed. The model inputs are based on prior medical research on comorbidities and evaluation by expert panels (Goldman et al. 2004). The functional form and inputs for these models are provided in Table A.2.2. These transition probabilities allow for the evolution of individual risk factor histories which are used to inform transitions across states in the state space. Medical comorbidities were treated as absorbing (i.e., irreversible) states; if one developed diabetes, for example, they had it until the end of their life (i.e., their probability of having diabetes in all successive transitions was set to one). Transition models for medical comorbidities were set up to estimate incident cases. Thus, for each medical comorbidity, respondents with the condition at baseline were excluded. For example, a respondent with diabetes at baseline but without a history of stroke would be excluded from the transition model for diabetes but included in the transition model for stroke. All states of BMI were considered transient (i.e., an individual could transition between being underweight, healthy weight, or overweight/obese) and no exclusions were made to the baseline sample to estimate the transition model. It is important to note that the transition matrices obtained from this process and therefore the relationships between covariates and the state space are held constant throughout the simulation. Changes in population composition and risk factor distribution are accounted for as described in Section 2.2.6.5.

2.2.6.5 Replenishing Cohorts

Individuals in the base population advance in age by one year with each one-year simulation cycle. A consequence of this process is that, after the first simulation cycle, the population no longer spans the full age distribution of interest; all 51-year-olds in the base population celebrate their 52nd birthday and there are no 51-year-olds to begin the

second simulation cycle. Thus, a mechanism that replenishes the synthetic population with new individuals aged 51 years old in each simulation cycle is needed to maintain population representativeness and to reflect demographic and epidemiological trends in the US. This gives rise to a challenge in that these trends are unobserved. A naïve approach to generate risk profiles for incoming cohorts would be to randomly sample from the base population and assign covariate values by sex and race/ethnicity. However, it is well established that trends in many of these risk factors—in addition to the risk of dementia—are changing over time (Wu et al. 2017b). This is especially true of educational attainment and obesity which have risen across successive birth cohorts (Prince et al. 2016). Thus, an approach that integrates these trends is warranted.

As with individuals in the base population, each new individual enters the model with a set of characteristics that, in aggregate, should reflect the joint distribution of the initial conditions of the base population as well as compositional and epidemiological trends observed in or applied to the population. However, as many of these entrants will be added in future years for which we have not observed data, it is necessary to either forecast trends in these risk factors or draw from external data sources. An additional consideration is that the population needs to maintain the correct size, structure, and representativeness while incorporating new entrants. To address these challenges, I combined HRS data with additional data sources described in Section 2.2.3 and Section 2.2.4 to produce synthetic replenishing cohorts.

In each simulation cycle c, a fixed number N of 51-year-olds enter the population with their M individual characteristics denoted by $y = (y_1, y_2, \dots, y_M)$. N is derived from population counts and forecasts provided by the Census (U.S. Census Bureau 2017,

2018) to reflect the size (i.e., count) and composition (i.e., sex and race/ethnicity) of the incoming cohorts of 51-year-olds in each simulation cycle for the years 2001-2050. *M* maps to the number of characteristics described in Section 2.2.2.7. These characteristics, which are correlated with each other, change over time. Denote by $F_t(y)$ the joint probability distribution of these *M* characteristics in year *t*. Let $F_m(y_m)$ represent the marginal distribution of the variable y_m. To model change in each successive replenishing cohort of 51-year-olds at time *t*, I examine the evolution of the marginal distribution for each of *M* characteristics, indicated by $\{F_{m,t}(y_m), m = 1, 2, \dots, M\}$. To do so, I model the change in the expected value of y_m and then apply an exogenous trend component to each replenishing cohort.

I begin by modeling historical trends in each of the *M* characteristics using observed data on successive cohorts of 51-year-olds from the BRFSS for the years 2002 to 2018. Hypertension was assessed bi-annually; all other covariates were assessed annually. I construct factors of change based on these historical trend components; these are applied to future replenishing cohorts for which we do not observe data (e.g., the cohort of 51-year-olds in the year 2030). The sampling frame of the BRFSS is sufficiently large to generate factors of change separately by sex (men, women) and race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, Hispanic) resulting in the estimation of trends for eight subgroups which may undergo changes in risk factors at different rates (Arroyo-Johnson and Mincey 2016; Pool et al. 2017). For example, among 51-year-olds in the BRFSS, between 2002 and 2018, the prevalence of diabetes among non-Hispanic white men increased from 6.4% to 11.2% whereas the prevalence of diabetes for non-Hispanic black men increased from 13.9% to 19.0% over

the same period; these trends correspond to overall changes of 75.0% and 36.7%, respectively, or average annual growth rates (AAGRs) of 4.4% and 2.2%. From the period 2002 to 2018, among 51-year-olds in the BRFSS, the proportion of non-Hispanic white women with a college degree increased from 36.3% to 46.8% (AAGR: 1.7%); the proportion of non-Hispanic black women with a college degree increased from 23.1% to 32.4% (AAGR: 2.4%).

However, calculating and applying the 17-year AAGR for each characteristic would require the assumption that the AAGR calculated during this period (2002 to 2018) would persist into the future for the years 2019-2050. This could overestimate the evolution of each characteristic if, for example, these characteristics underwent substantial change during this period but the rate of change levelled off (e.g., if trends in educational attainment increase at a decreasing rate). For example, the proportion of non-Hispanic white women with a college degree changed from 36.3% in 2002, to 37.8% in 2010, to 46.8% in 2018; the corresponding AAGRs for the trend from 2002 to 2010 and 2010 to 2018 are 0.45% and 2.65%, respectively. As noted earlier, the AAGR from 2002 to 2018 was 1.7%. The latter provides valuable information on the AAGR over the 17-year period whereas the former indicates that the proportion of non-Hispanic white women attaining a college degree is increasing at an increasing rate. Thus, rather than apply the AAGR for the period 2002 to 2018, I estimated the AAGR for four five-year periods: 2002-2006, 2006-2010, 2010-2014, and 2014-2018. This allowed for the identification of a constancy or leveling off effect while incorporating a degree of uncertainty about future trends.

The changes in the five-year AAGRs themselves were used to inform the rates at which each characteristic would change over time for unobserved years. That is, in addition to accounting for changes in risk profiles of 51-year-olds as it pertains to risk factors, the rates at which these changes occurred in five-year periods were embedded within the simulation.

The factors of change (or trend components) were implemented as follows. Observed trends in the BRFSS were applied for the years 2001-2018 and combined with actual population counts obtained from the Census for the years corresponding to the simulation. A similar procedure was implemented for the simulation cycles corresponding to the years 2019 to 2050. The population counts by sex and race/ethnicity were obtained from Census projections. The risk factor distributions were obtained by shifting the distribution of these measures in accordance with the AAGRs in five-year intervals. Table A.2.3 provides sample characteristics for select replenishing cohorts. Overall, upward trends were observed for educatoinal attainment and obesity while a downward trend were observed for smoking. For simplicity, and due to the low prevalence of dementia below the ages of 65, it was assumed that all 51-year-olds entered the model without cognitive impairment.

2.2.6.6 Simulation

The core of the microsimulation is to evolve the population through time to represent in each annual cycle the demographic and epidemiological structure of the population and the prevalence of outcomes in the state space {no cognitive impairment, cognitive impairment without dementia, undiagnosed dementia, diagnosed dementia, death}. This is achieved by aging the population forward in time, replenishing the younger ages in the

population with new cohorts of individuals whose joint covariate distributions approximate trends in dementia risk factors, and evolving the population's covariate profiles.

The simulation engine operates within a discrete-time framework and it is assumed that a maximum of one transition can occur within each simulation cycle. Two steps were taken to improve computational efficiency. First, rather than simulate the population in its entirety, I took random 5% samples from the base population and each replenishing cohort. This process was implemented for each of 500 runs of the simulation (with replacement as noted in Section 2.2.6.10). Second, the simulation was conducted in blocks of 1,000 individuals to mitigate any potential issues related to memory allocation. For example, rather than simulate and store output for 10 million individuals simultaneously, the model operates by simulating blocks of 1,000 individuals, storing the output, deallocating the memory, and repeating this process 10,000 times—each with a different sample of 1,000 individuals. With the base population, transition probabilities, and replenishing cohorts defined, I implemented a 3-step algorithm for the simulation process.

1. An individual either starts in the base population or enters through a replenishing cohort with a set of covariates that define their risk profile and current state.
2. In each cycle, an individual's probability of transitioning between states is calculated as a function of their demographic, lifestyle, and health characteristics as well as their previously occupied state(s). These probabilities are derived from the observed transition matrices obtained from the HRS and are assumed to remain constant.

3. A random variable is drawn from a Bernoulli (if only two states) or categorical (if more than two states) distribution shaped by the transition probabilities in Step 2; this determines the state to which an individual will transition (or remain).

Steps 1-3 were then repeated for 50 simulation cycles. This entire process was repeated 500 times as described in Section 2.2.6.10. Characteristics of each individual are tracked through each simulation cycle making it possible to know in which year an individual transitioned between states as well as the duration of time spent in each state. The dynamics of the microsimulation process are depicted in Figure 2.2.

2.2.6.7 Output

The simulation provides estimated population counts for the state space over 50 annual cycles. That is, for each calendar year from 2000-2050, the model outputs the number of adults without cognitive impairment, with mild cognitive impairment, with undiagnosed dementia, with diagnosed dementia, and the number of deaths. The output also tracks the number of years spent in each state. All outputs can be broken down by demographic, lifestyle, and health characteristics.

2.2.6.8 Internal Validation

The proportion of adults in each state of the state space was validated internally and retrospectively for the period 2000-2014. This was done by comparing the actual prevalence estimates in the HRS and its linkages to the simulated estimates. The transition probabilities used in the simulation relied on data from survey years 2000-2010; thus, in addition to reporting modeling error, any discrepancy between the

simulated and actual rates in the HRS for the survey years 2012 and 2014 provides a mechanism for interval validation.

2.2.6.9 External Validation

To validate all-cause mortality, I compared age-specific mortality rates derived from the simulation output to period estimates provided by the Human Mortality Database for the year 2010. I also validated the size of the population aged 65 years or over by comparing results obtained from the simulation for the years 2000 to 2050 to Census projections for the same age group and period.

2.2.6.10 Uncertainty

Sources of uncertainty were introduced into the simulation process to obtain probabilistic estimates. As noted in Section 2.2.6.4, transition models were estimated 500 times using a bootstrap approach. Draws from the base population were made—with replacement—to account for the joint distribution of population characteristics. The transition parameters were then estimated each of these 500 times to generate uncertainty bounds for the parameter distribution. An additional source of uncertainty for the transitions occurred at the simulation stage; an individual's transition between time t and $t + 1$ was determined by sampling a random variable from a Bernoulli (if only two states) or categorical (if more than two states) distribution based on their probability of transitioning into each state of the state space. Uncertainty was also introduced into the replenishing cohorts by assigning each new entrant's risk profile from a bootstrapped sample of the joint distribution of population characteristics from the base population, with modifications for trends as described in Section 2.2.6.5. Five hundred runs of the

complete simulation process were completed to generate a range of outcomes taking into account these sources of uncertainty.

2.2.6.11 Assumptions

The Dementia Population Cost Model is a complex system, both in terms of computation and parameter identifiability, that patches together a variety of data sources and identification schemes. As such, assumptions about the processes and mechanisms underlying each phase of the system were introduced. Table A.2.4 presents the underlying assumptions for the Dementia Population Cost Model.

2.3 Results

Baseline characteristics of the sample prior to and after imputation are shown in Table 2.1. Weighted and unweighted sample characteristics are shown in Table 2.2. The 18,277 HRS respondents at baseline mapped to 34,896,000 individuals after applying the modified sampling weights. The sample weighted proportion of individuals with cognitive impairment but free of dementia (14.5), undiagnosed dementia (4.43), and diagnosed dementia (4.09) at baseline shown in Table 2.2 were higher than the unweighted sample proportions. This was partially attributable to upweighting of younger adults who have lower prevalence rates of cognitive impairment.

A comparison between the weighted HRS population counts and estimates obtained from the US census are shown for the year 2000 by age and sex in Figure A.2.1. There are clear discrepancies in the youngest and oldest age bands; the weighted HRS undercounts the number of adults between ages 51 and 54 by a factor of nearly 1.8 (Figure A.2.1, Panel B) when using the HRS-provided sampling weights. This may be

due to the complex sampling design of the HRS which, when sample weights are applied, map to the non-institutionalized and nursing home population over the age of 50 whereas the US Census population counts represent all community-dwelling and institutionalized populations. This large discrepancy helps validate the additional procedure used in applying the sample weights to obtain a base population more representative of the population in the year 2000 in accordance with the Census.

Forecasted population counts (in millions) by cognitive status from 2000 to 2050 are presented in Figure 2.3. These estimates and their corresponding ranges for the years 2010, 2030, and 2050 are shown in Table 2.3 (in thousands) overall and for men and women. The model, which takes into account demographic and epidemiological changes in the population, forecasts an increase in the number of adults aged 65 years and over with some form of cognitive impairment (CIND, undiagnosed dementia, diagnosed dementia) from 11.21 million (range from 500 simulations: 9.69, 13.71) in 2010 to 18.82 (range: 16.34, 22.39) in 2030 and 27.45 (range: 23.33, 32.27) million in 2050 which correspond to respective increases of 1.68 (range: 1.46, 1.99) in 2030 and 2.45 (range: 2.08, 2.88) in 2050 relative to 2010. The number of adults with undiagnosed or diagnosed dementia (hereafter referred to strictly as dementia) is estimated to increase from 6.32 million (range: 5.33, 7.72) in 2010 to 10.18 (range: 9.03, 12.21) in 2030 and 16.29 million (range: 14.02, 19.05) in 2050, corresponding to a more than two-fold increase from 2010 to 2050.

The crude and age-standardized prevalence rates of dementia in the population aged 65 years or older are shown in Figure 2.4, with crude rates shown in Table 2.4. The crude rates are influenced by changes in the age structure of the population over time as well as

compositional (e.g., proportion of racial/ethnic minorities, distribution of educational attainment) and epidemiological (e.g., risk of diabetes) changes in the population. The age-standardized rates account for changes in the age structure of the population by mathematically adjusting the age structures so they are comparable. In this context, the population age structures for the years 2001 to 2050 were standardized to the year 2000 population age structure. Thus, to a degree, forecasted trends in dementia prevalence across years can be examined in the absence of population aging. The crude prevalence of dementia in 2010 was estimated at 16.8%, 15.0% in 2030, and 20.2% in 2050. Comparatively, the age-standardized prevalence of dementia was estimated at 16.7% in 2010, 15% in 2030, and 18% in 2050. The relative stability of the age-standardized compared with the crude prevalence of dementia from 2000 to 2050 indicates that the increasing prevalence is largely attributable to population aging.

Much of the increase in the number of adults with dementia occurs in the older age groups (Figure 2.5, Table 2.3 Panel A). Overall, the number of adults aged 65-74 years with dementia is expected to increase from 1.65 million (range: 1.35, 1.99) in 2010 to 2.75 (range: 2.66, 3.59) in 2030 and 3.93 million (range: 3.31, 4.46) in 2050, corresponding to a 20-year increase from 2010 of 1.67 (range:1.61, 2.18) and a 40-year increase from 2010 of 2.38 (range: 2.09, 2.7). The number of adults aged 75-84 years with dementia is expected to increase from their levels in 2010 by a factor of 1.66 (range: 1.42, 1.97) in 2030 and by 2.32 (range: 1.99, 2.76) in 2050. For adults aged 85 years or older, 20- and 40-year increases from 2010 of 1.51 (range: 1.29, 1.72) and 3.05 (range: 2.67, 3.58), respectively, are expected. The model also indicates that the number of women aged 65 years or older with dementia in 2050 (9.79; range: 8.55, 11.53) will

exceed the number of men with dementia in the same year (6.49; range: 5.45, 7.52). This is driven primarily by women aged 85 years or older, which is consistent with the extant literature (i.e., there is observed divergence in the risk of dementia among men and women at around age 80). This may also be driven by differences in the size of the 85+years age group among men and women, with women having longer average life expectancy and thus being more represented at the oldest ages when dementia is most likely to occur.

Figure 2.6 shows considerable differences in the expected number of adults with dementia by sex and racial/ethnic groups in the coming decades. All subgroups exhibit some degree of increase between 2000 and 2050. The estimates obtained from the model suggest Hispanic men and women will experience the largest and fastest increases in dementia, perhaps due to these subgroups being the fastest growing demographic in the US. With the exception of non-Hispanic black adults, all subgroups show a steady or rapid decline in the absolute number of adults with dementia. Non-Hispanic black men and women show a relatively steady trajectory (through about year 2020 for women, and year 2025 for men) until increasing in conjunction with their counterparts.

Figure 2.7 displays the forecasted total number of adults aged 65 years or over with dementia from 2000 to 2050 together with the actual counts from the weighted HRS sample for the years 2000 to 2014. The actual state occupancies from the HRS for all years—including 2012 and 2014, which were not included in the estimation sample—fall within the bounds of the forecast indicating the model appropriately captured the aging dynamics within the HRS and, more generally, within the population. The model forecasts increasing counts of adults over the age of 65 with some level of cognitive

impairment from 2000 to 2050. Forecast uncertainty, as indicated by the gradual widening of the uncertainty bands, increases over time.

To validate the model externally, I compared the age-specific mean mortality rates obtained from the simulation with period mortality rates obtained from the HMD for the years 2000 to 2010. These results are shown in Figure 2.8. The simulated mortality rates show excellent concordance with those obtained from the HMD. Figure 2.9, which compares the forecasted total population size for adults aged 65 years or over for the years 2000 to 2050 derived from the simulation and Census projections, provides further evidence of model validation. As expected, the simulation undercounts Census projections to a small degree; this is at least partially attributable to the fact that mortality trends were not directly modeled. Thus, projections that account for reduced mortality trends will produce high population counts.

2.4 Discussion

In this chapter, I introduced a probabilistic approach to forecasting the prevalence of mild cognitive impairment and dementia using a microsimulation framework. Forecasts obtained from the Dementia Population Cost Model suggest that the number of adults over the age of 50 with dementia will more than double over the next 30 years due to population aging and trends in the underlying risk of dementia. Results obtained from the model indicate that Hispanic men and women will face the largest relative increase in dementia which could be attributable to their higher dementia risk relative non-Hispanic white adults and their status as the fastest growing racial/ethnic segment of US adults aged 65 years or over, with a projected nearly tripling of representation in the United States from 8% in 2014 to 22% by 2060 (Robinson 2016).

The Dementia Population Cost Model is unique among existing models for dementia projections that use extrapolation (Matthews et al. 2019; Prince et al. 2016; Prince et al. 2013; Prince and Jackson 2009) or macrosimulation approaches (Brookmeyer et al. 1998; Hebert 2000). In its simplest form, extrapolation methods project the future number of cases of dementia by estimating strata-specific dementia prevalence and multiplying by future population projections. In fact, even the projections published by Alzheimer's Disease International (Prince et al. 2016; Sloane et al. 2002) use this approach which assumes that age- and sex-specific prevalence of dementia and its risk factors will not vary over time, attributing any future changes in its prevalence to population aging alone. Studies using a macrosimulation-based approach tend to incorporate more dynamic information in their models, such as strata-specific incidence rates of dementia, but often apply these rates under the assumption that the underlying risk factors are stable over time.

The Dementia Population Cost Model takes a probabilistic approach to forecasting the prevalence of different levels of severity of dementia in a microsimulation framework that directly accounts for changes in demographic and epidemiological forces. In addition, despite the extant literature focusing on the prevalence of dementia at age 65 and above, the current approach simulates the progression of medical comorbidities, such as diabetes and hypertension, at mid-life which are believed to be more indicative of dementia risk relative to their measurement in later life (Barnes and Yaffe 2011a; Exalto et al. 2014; Xu et al. 2009). Further, it is the only model that links Medicare claims information to a large, nationally representative sample of older adults to forecast the number of adults with undiagnosed and diagnosed dementia. A growing body of work

has reported that claims-based diagnoses of dementia may undercount the true prevalence among racial/ethnic subgroups (Chen et al. 2019) while survey-based measures of dementia may be criticized for their lack of a clinical diagnostic procedure. Incorporating states of undiagnosed and diagnosed dementia provides more information about the true prevalence of dementia in the population which allows for a better assessment of the overall burden. In addition, this information can advance our understanding of the future screening needs among population strata characterized by sociodemographic characteristics and medical comorbidities.

The current study accounted for risk factor trends that showed both improvements (e.g., increasing educational attainment, reduced rates of smoking) and regression (e.g., increased rates of diabetes) in the underlying risk of dementia through the inclusion of replenishing cohorts. Thus, a priori, it was not clear how the current model would compare to prior estimates obtained via extrapolation. Accounting for changes in the underlying risk of dementia produced previously undocumented trends in its future burden. Despite the relatively stable prevalence rates for dementia shown in Table 2.4, the apparent decline in prevalence shown in Figure 2.4 is consistent with a growing body of working documenting a decline in age-specific prevalence of dementia. This is also consistent with findings from Chapter 1 which indicated that improvements in population-level risk factors produce lower incidence rates of dementia. However, Scenario 7 from Chapter 1, which additionally accounted for population-level increases in diabetes and obesity, was indicative of a rebound in the risk of dementia. The observed decline in the age-standardized prevalence of dementia shown in the current study could stem from increases in educational attainment embedded within the simulation; the

apparent rebound in age-standardized prevalence shown in Figure 2.4 Panel B could be indicative of population-level improvements in educational attainment reaching a ceiling effect at which point the increasing risks of comorbid conditions begin to offset the reductions attained through increasing education. .In other words, it is plausible that rising levels of comorbidities are balanced by increasing levels of educational attainment to a certain inflection point, at which improvements in educational attainment stagnate while the prevalence of comorbid conditions continues to rise. To my knowledge, this is the only forecast of dementia that successfully captures declining trends in prevalence and shows a rebound effect.

Comparisons between estimates obtained from the current model and those used in prior studies should be examined carefully and in full view of their different approaches, data sources, inputs, and time periods under study. Prior studies using extrapolation methods are static in that they don't account for changes in underlying dementia risk. If trends in the underlying risk of dementia increase over time, then these static methods should be an underestimate (all things being equal). Conversely, if trends in the underlying risk of dementia decrease over time, extrapolation methods should be an overestimate relative to methods accounting for such changes (all things being equal). With these thoughts in mind, Table A.2.5 provides a summary of existing studies estimating the future burden of dementia and Figure 2.10 presents a comparison of dementia prevalence obtained from the Dementia Population Cost Model with previous studies. The estimates obtained from the current study exceed prior estimates although it is important to note much of the prior literature focused exclusively on Alzheimer's disease which is one of several causes of dementia. In addition, most prior work has

focused exclusively on survey- or claims-based data; using only one would undercount the total burden of dementia.

Among prior studies, the results from the Dementia Population Cost Model were most similar to those obtained from Zissimopoulos and colleagues (Zissimopoulos et al. 2018) which used the well-known Future Elderly Model. To date, Zissimopoulos et. al. produced one of the highest forecasts for the size of the population with dementia. However, compared to the current study, their estimates fall between the mean estimation from the simulation and the lower bound from 500 simulations. The study by Zissimopoulos classified adults with dementia using only survey-based measures; thus, their study is likely to undercount the absolute number of adults with dementia, especially among adults whose cognitive impairment is not yet recognized by their healthcare system; which itself may be a reflection of their access to healthcare, how they utilize services, and their interactions with healthcare systems.

The estimates obtained from the Dementia Population Cost Model support the inclusion of both undiagnosed and diagnosed dementia to mitigate undercounting of the population with dementia. This is especially important considering the recent literature indicating the varying degrees of underreporting of diagnosed dementia in administrative databases (Chen et al. 2019). This is suggestive that prior work using only claims-based data may undercount the future burden among minority subgroups. In addition, based on the comparison of results to Zissimopoulos and colleagues (Zissimopoulos et al. 2018), it indicates that including only survey-based measures may also result in undercounting.

In its current iteration, the model could be used for counterfactual simulation to assess dynamic relationship between the future prevalence of dementia and changes in

population composition as well as demographic and epidemiological forces. However, the model could be expanded to include cost estimates, how they vary among demographic strata and comorbidity profiles, and how they evolve over time with respect to type of impairment and time since diagnosis. It would be useful to understand how these costs vary by type of insurance as well as shifts in insurance coverage over time. Future iterations of the model could also account for changes in diagnostic procedures which may artificially inflate the prevalence of dementia (e.g., through biomarker-based clinical definitions). Embedding this type of information within the model would add to its utility as a mechanism for counterfactual simulation to assess the impact of policy changes on the future economic burden of dementia, particularly with respect to evolving public and private insurance systems.

I acknowledge that the current iteration of the model requires refinements. First, while I am able to account for changes in risk factor distributions, I assumed their relationships are constant. That is, I incorporate distributional changes over time but do not account for relational changes that may occur. This implicitly allows for individual mortality risk to change, for instance, in response to increasing educational attainment. However, no explicit assumptions are made about how population-level mortality rates change over time. Second, although I incorporated what are considered to be key clinical risk factors for cognitive impairment, for simplicity, it was assumed that these measures were strictly additive. Not accounting for multiplicative effects on cognitive status and mortality across these risk factors could under or overestimate dementia prevalence depending on their synergistic effect on both cognitive status and mortality. Third, strong assumptions were made about future incoming cohorts of 51-year-olds. The population

count and composition of incoming 51-year-olds in each simulation cycle were derived from the US Census Bureau. Time trends in risk factors for these incoming cohorts were extrapolated from a large, nationally representative annual survey of adults in the US. It is possible that one or more of these risk factors could increase or decrease at a faster or slower rate than the average annual growth rate observed in recent decades. Fourth, the model does not account for the use of medications which some individuals may be prescribed to manage their comorbidities nor does it account for future advancements in medical science or technology. As the medical community works to improve health outcomes for older adults, it is possible that individuals will be living longer with more comorbidities. This could result in an underestimation of cognitive impairment and an overestimation of mortality. Fifth, I assumed the combined person-level and nursing home sampling weights provided by the HRS sufficiently map the HRS sample to the US population of adults over the age of 50. There are clear discrepancies between the population counts as indicated by Figure A.2.1 but most age groups showed good concordance between the two sources. These discrepancies may be attributed to differences in the samples; the HRS is considered nationally representative of community-dwelling and nursing home populations whereas the Census includes community-dwelling and institutionalized populations. The main group quarters included in the Census that may not be accounted for in the HRS include, for example, correctional institutions, military housing, other institutional facilities, and other noninstitutional facilities. The population counts obtained from the simulation for adults aged 65 years or over for the period 2000 to 2050 undercount but closely resemble those obtained from Census projections for the same age group and period. The simulation

included annual population projections for 51-year-olds from the Census but did not directly model changes in mortality. Thus, the slightly higher mortality rates used in the simulation resulted in a higher proportion of adults dying in each cycle and an overall lower population count. The undercounting in earlier years is due to the weighting procedure which proportionally upweighted respondents in the HRS based on the HRS-provided sampling weights to approximate the Census counts without overrepresenting any subgroup (defined by sex and race/ethnicity). Higher concordance with Census counts and forecasts would likely yield higher counts of dementia; thus, to an extent, the estimates reported here may serve as a lower bound relative to what would have been observed with a higher population count on part with those reported and projected by the Census. Finally, it is important to note that the estimates obtained in the current study are not projections of the impending dementia burden but are instead probabilistic forecasts of what may occur on the basis of these caveats and the assumptions described in Table A.2.4.

Despite these caveats, it is evident from the current model and prior work using alternative methodologies that the absolute number of adults with dementia will increase in the US in response to population aging which will likely outweigh the potential reductions due to improvements in underlying risk factors. This increase will occur disproportionally among minority subgroups who will comprise an increasing share of the population over age 65 and who face greater prevalence of dementia, underscoring the importance among social and health care systems of monitoring and evaluating minority health and developing strategies to effectively deliver health care services that

meet the social, cultural, and linguistic needs of patients and their caregivers(Betancourt, Green and Carrillo 2002).

In conclusion, the Dementia Population Cost Model shows that population aging is likely to result in an increase in the number of adults with dementia despite current and forecasted trends revealing improvements in the underlying risk of dementia. These estimates carry with them important social and economic implications that can be used to inform healthcare and policy planning efforts. Existing projections that use extrapolation or macrosimulation methods show how the scale of dementia will change in response to population aging but typically assume that a constancy of strata-specific dementia prevalence and its underlying risk factors. The approach of the current study is more informative as it allows for a detailed examination of the future burden of dementia in the context of demographic and epidemiological forces in the presence of uncertainty. Counterfactual scenarios can be examined using the Dementia Population Cost Model to understand how changes in population-level risk factors may delay or prevent dementia onset. In addition, future iterations that incorporate evolution of payer mix and economic information will be useful for understanding how the public and private costs of dementia may change as a function of insurance coverage.

2.5 Tables

Table 2.1 Unweighted sample characteristics prior to and after imputation, HRS 2000

Characteristic	Before imputation		After imputation	
	n	%	n	%
Age				
51-64	7,825	42.81	7,825	42.81
65-74	5,546	30.34	5,546	30.34
75-84	3,617	19.79	3,617	19.79
85+	1,289	7.05	1,289	7.05
Sex				
Male	7,788	42.61	7,788	42.61
Female	10,489	57.39	10,489	57.39
Race/ethnicity				
Non-Hispanic white	14,014	76.68	14,014	76.68
Non-Hispanic black	2,494	13.65	2,494	13.65
Non-Hispanic other	357	1.95	357	1.95
Hispanic	1,412	7.73	1,412	7.73
Education				
< HS or GED	5,849	32	5,849	32
HS	5,736	31.38	5,736	31.38
Some college	3,502	19.16	3,502	19.16
College and above	3,190	17.45	3,190	17.45
Smoking status				
Never	7,478	40.91	7,548	41.3
Former	7,982	43.67	8,073	44.17
Active	2,653	14.52	2,656	14.53
Missing	164	0.9	0	0
BMI				
Underweight	384	2.1	384	2.1
Healthy weight	6,227	34.07	6,248	34.19
Overweight or obese	11,413	62.44	11,645	63.71
Missing	253	1.38	0	0
History of diabetes				
No	15,549	85.07	15,569	85.18
Yes	2,708	14.82	2,708	14.82
Missing	20	0.11	0	0
History of hypertension				
No	9,635	52.72	9,640	52.74
Yes	8,624	47.18	8,637	47.26
Missing	18	0.1	0	0
History of heart disease				
No	14,248	77.96	14,258	78.01
Yes	4,017	21.98	4,019	21.99
Missing	12	0.07	0	0
History of stroke				
No	16,780	91.81	16,791	91.87
Yes	1,486	8.13	1,486	8.13
Missing	11	0.06	0	0

Notes. BMI, body mass index; GED, general education development test; HS, high school.

Table 2.2 Weighted and unweighted sample characteristics of the base population, HRS 2000

Characteristic	Unweighted		Weighted	
	n	%	n	%
Age				
51-64	7,825	42.81	31,424,081	48.5
65-74	5,546	30.34	17,686,934	27.3
75-84	3,617	19.79	11,926,593	18.41
85+	1,289	7.05	3,751,888	5.79
Sex				
Male	7,788	42.61	28,732,728	44.35
Female	10,489	57.39	36,056,768	55.65
Race/ethnicity				
Non-Hispanic white	14,014	76.68	53,327,647	82.31
Non-Hispanic black	2,494	13.65	5,951,789	9.19
Non-Hispanic other	357	1.95	1,379,460	2.13
Hispanic	1,412	7.73	4,130,600	6.38
Education				
< HS or GED	5,849	32	18,380,982	28.37
HS	5,736	31.38	20,437,276	31.54
Some college	3,502	19.16	13,195,166	20.37
College and above	3,190	17.45	12,776,072	19.72
Smoking status				
Never	7,548	41.3	26,414,134	40.77
Former	8,073	44.17	28,372,275	43.79
Active	2,656	14.53	10,003,087	15.44
BMI				
Underweight	384	2.1	1,289,636	1.99
Healthy weight	6,248	34.19	22,177,477	34.23
Overweight or obese	11,645	63.71	41,322,383	63.78
History of diabetes				
No	15,569	85.18	56,052,774	86.52
Yes	2,708	14.82	8,736,722	13.48
History of hypertension				
No	9,640	52.74	35,862,888	55.35
Yes	8,637	47.26	28,926,608	44.65
History of heart disease				
No	14,258	78.01	51,411,695	79.35
Yes	4,019	21.99	13,377,801	20.65
History of stroke				
No	16,791	91.87	59,876,353	92.42
Yes	1,486	8.13	4,913,143	7.58
State space				
No cognitive impairment	13,563	74.21	49,877,491	76.98
CIND	2,931	16.04	9,392,055	14.5
Undiagnosed dementia	956	5.23	2,867,184	4.43
Diagnosed dementia	827	4.52	2,652,766	4.09

Notes. BMI, body mass index; CIND, cognitive impairment without dementia; GED, general education development test; HS, high school.

Table 2.3 Forecasted number of adults without cognitive impairment or with CIND, undiagnosed dementia, or diagnosed dementia by age group overall (A) and for men (B) and women (C), 2010, 2030, 2050
(A)

	Forecasted population counts (in thousands), Overall		
Age group and state	2010	2030	2050
65+ years			
Healthy	26257 (23893, 32336)	48827 (43540, 57747)	52988 (46179, 62680)
CIND	4896 (4360, 5992)	8641 (7310, 10193)	11160 (9311, 13218)
Undiagnosed dementia	1659 (1412, 2051)	2580 (2358, 3333)	4118 (3381, 4706)
Diagnosed dementia	4657 (3919, 5669)	7603 (6676, 8873)	12172 (10642, 14343)
65-74 years			
Healthy	16325 (14924, 20385)	29128 (25723, 34164)	27775 (24179, 33269)
CIND	2105 (1815, 2689)	3903 (3330, 4559)	4188 (3364, 5039)
Undiagnosed dementia	789 (663, 954)	1354 (1316, 1883)	2068 (1744, 2275)
Diagnosed dementia	861 (691, 1043)	1395 (1343, 1715)	1858 (1563, 2187)
74-85 years			
Healthy	7684 (6987, 9580)	15645 (14213, 18894)	17154 (14699, 20038)
CIND	1945 (1730, 2261)	3365 (2813, 3968)	4151 (3590, 4798)
Undiagnosed dementia	611 (458, 769)	919 (804, 1150)	1413 (1133, 1609)
Diagnosed dementia	1962 (1708, 2312)	3365 (2851, 3913)	4557 (4003, 5495)
85+ years			
Healthy	2247 (1983, 2371)	4055 (3603, 4689)	8059 (7300, 9373)
CIND	847 (814, 1042)	1373 (1167, 1676)	2820 (2356, 3380)
Undiagnosed dementia	259 (219, 328)	308 (237, 366)	637 (505, 822)
Diagnosed dementia	1834 (1520, 2315)	2844 (2483, 3244)	5757 (5077, 6661)

Notes. CIND, cognitive impairment without dementia. Data in parentheses are range from 500 simulations.

(B)

Age group and state	Forecasted population counts (in thousands), Men		
	2010	2030	2050
65+ years			
Healthy	11517 (10743, 14211)	22633 (20159, 26843)	24857 (21579, 29338)
CIND	2105 (1986, 2722)	4065 (3484, 4677)	5197 (4439, 6240)
Undiagnosed dementia	821 (668, 1006)	1170 (1097, 1644)	2038 (1622, 2208)
Diagnosed dementia	1598 (1393, 1952)	2714 (2419, 3304)	4458 (3855, 5316)
65-74 years			
Healthy	7528 (7031, 9295)	13945 (12208, 16339)	13552 (11787, 16062)
CIND	1018 (920, 1360)	1870 (1690, 2205)	1968 (1707, 2507)
Undiagnosed dementia	440 (366, 491)	628 (619, 922)	1060 (874, 1094)
Diagnosed dementia	343 (295, 463)	573 (547, 730)	774 (645, 950)
74-85 years			
Healthy	3165 (3070, 4115)	7086 (6463, 8647)	7854 (6750, 9245)
CIND	792 (752, 956)	1560 (1282, 1746)	2030 (1675, 2248)
Undiagnosed dementia	280 (193, 392)	422 (348, 581)	693 (534, 727)
Diagnosed dementia	812 (683, 846)	1257 (1102, 1528)	1833 (1533, 2259)
85+ years			
Healthy	824 (641, 801)	1601 (1488, 1857)	3451 (3042, 4031)
CIND	295 (214, 406)	635 (511, 726)	1199 (1058, 1485)
Undiagnosed dementia	101 (89, 123)	121 (109, 141)	285 (214, 387)
Diagnosed dementia	442 (415, 644)	884 (770, 1045)	1851 (1676, 2107)

Notes. CIND, cognitive impairment without dementia. Data in parentheses are range from 500 simulations.

(C)

Age group and state	Forecasted population counts (in thousands), Women		
	2010	2030	2050
65+ years			
Healthy	14740 (13150, 18125)	26194 (23381, 30904)	28132 (24600, 33342)
CIND	2791 (2373, 3270)	4576 (3826, 5516)	5962 (4871, 6977)
Undiagnosed dementia	838 (744, 1046)	1411 (1260, 1689)	2080 (1760, 2498)
Diagnosed dementia	3059 (2526, 3717)	4889 (4257, 5569)	7715 (6788, 9027)
65-74 years			
Healthy	8797 (7893, 11090)	15183 (13515, 17825)	14223 (12392, 17207)
CIND	1087 (895, 1329)	2033 (1640, 2354)	2220 (1657, 2532)
Undiagnosed dementia	349 (297, 463)	726 (687, 961)	1008 (870, 1181)
Diagnosed dementia	518 (396, 580)	822 (795, 985)	1084 (917, 1237)
74-85 years			
Healthy	4519 (3916, 5465)	8558 (7751, 10247)	9301 (7950, 10793)
CIND	1152 (978, 1305)	1805 (1531, 2222)	2121 (1916, 2550)
Undiagnosed dementia	331 (265, 377)	497 (457, 569)	720 (599, 882)
Diagnosed dementia	1150 (1026, 1467)	2107 (1749, 2385)	2725 (2470, 3235)
85+ years			
Healthy	1424 (1341, 1570)	2453 (2115, 2832)	4608 (4259, 5342)
CIND	551 (500, 637)	738 (655, 939)	1621 (1299, 1895)
Undiagnosed dementia	158 (133, 206)	188 (117, 259)	352 (291, 435)
Diagnosed dementia	1392 (1105, 1670)	1960 (1713, 2199)	3906 (3401, 4555)

Notes. CIND, cognitive impairment without dementia. Data in parentheses are range from 500 simulations.

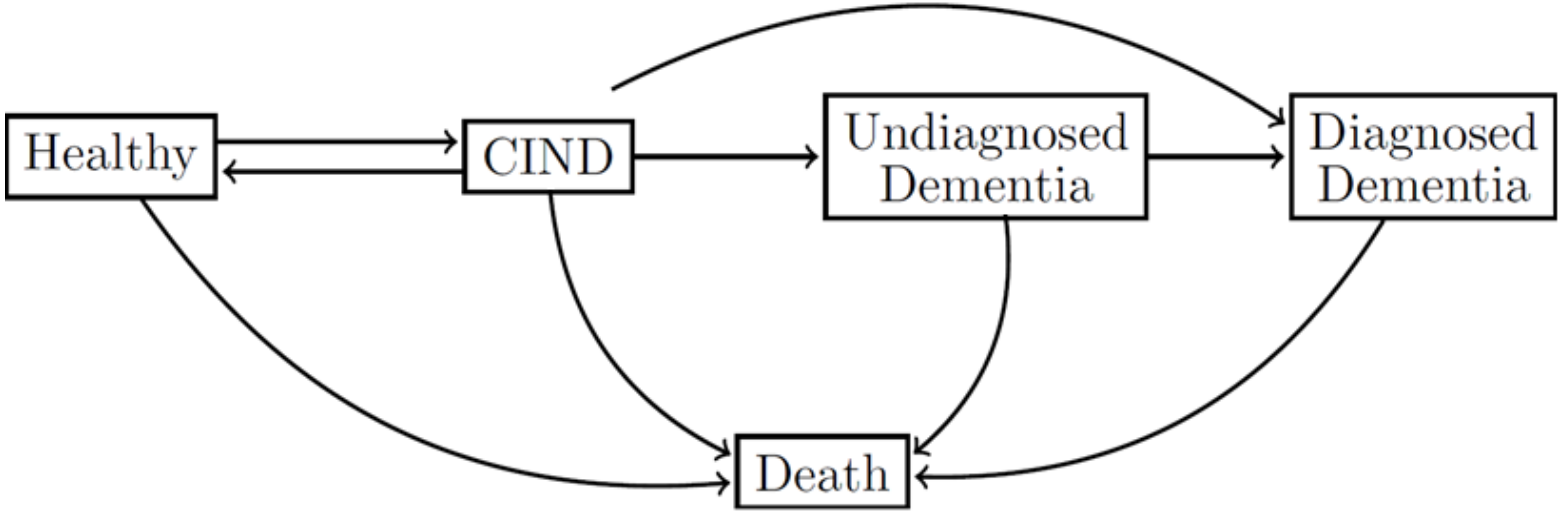
Table 2.4 Forecasted crude prevalence of adults without cognitive impairment or with CIND, undiagnosed dementia, or diagnosed dementia by age group overall and for men and women 2010, 2030, 2050

	Crude prevalence (%)								
	2010			2030			2050		
Age group and state	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
65+ years									
Healthy	70.1	71.8	68.8	72.2	74.0	70.7	65.9	68.0	64.1
CIND	13.1	13.1	13.0	12.8	13.3	12.3	13.9	14.2	13.6
Undiagnosed dementia	4.4	5.1	3.9	3.8	3.8	3.8	5.1	5.6	4.7
Diagnosed dementia	12.4	10.0	14.3	11.2	8.9	13.2	15.1	12.2	17.6
65-74 years									
Healthy	81.3	80.7	81.8	81.4	82.0	80.9	77.4	78.1	76.7
CIND	10.5	10.9	10.1	10.9	11.0	10.8	11.7	11.3	12.0
Undiagnosed dementia	3.9	4.7	3.2	3.8	3.7	3.9	5.8	6.1	5.4
Diagnosed dementia	4.3	3.7	4.8	3.9	3.4	4.4	5.2	4.5	5.8
74-85 years									
Healthy	63.0	62.7	63.2	67.2	68.6	66.0	62.9	63.3	62.6
CIND	15.9	15.7	16.1	14.4	15.1	13.9	15.2	16.4	14.3
Undiagnosed dementia	5.0	5.5	4.6	3.9	4.1	3.8	5.2	5.6	4.8
Diagnosed dementia	16.1	16.1	16.1	14.4	12.2	16.2	16.7	14.8	18.3
85+ years									
Healthy	43.3	49.6	40.4	47.3	49.4	45.9	46.7	50.9	43.9
CIND	16.3	17.7	15.6	16.0	19.6	13.8	16.3	17.7	15.5
Undiagnosed dementia	5.0	6.1	4.5	3.6	3.7	3.5	3.7	4.2	3.4
Diagnosed dementia	35.4	26.6	39.5	33.1	27.3	36.7	33.3	27.3	37.2

Notes. Crude prevalence shown. CIND, cognitive impairment without dementia. Subgroup totals may not aggregate to 100 due to rounding.

2.6 Figures

Figure 2.1 Five-state model for survival and cognitive status



Notes. CIND, cognitive impairment without dementia. Age and covariate transitions are not depicted for simplicity.

Figure 2.2 Dementia Population Cost Model Framework

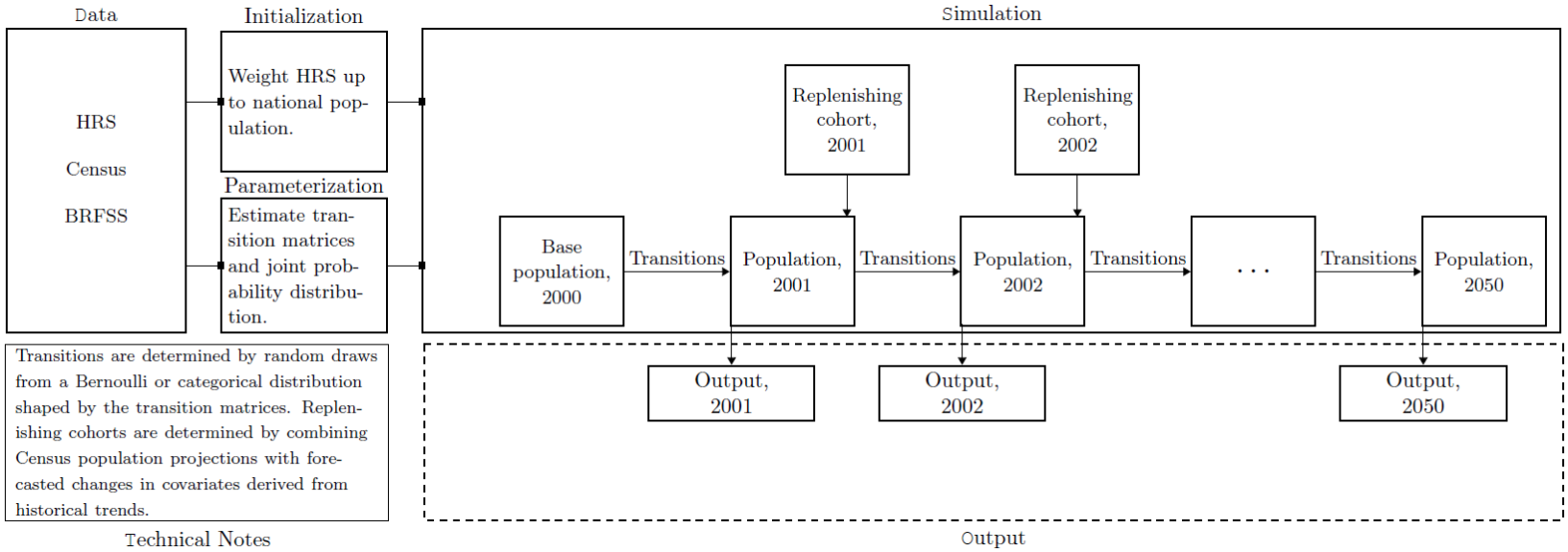
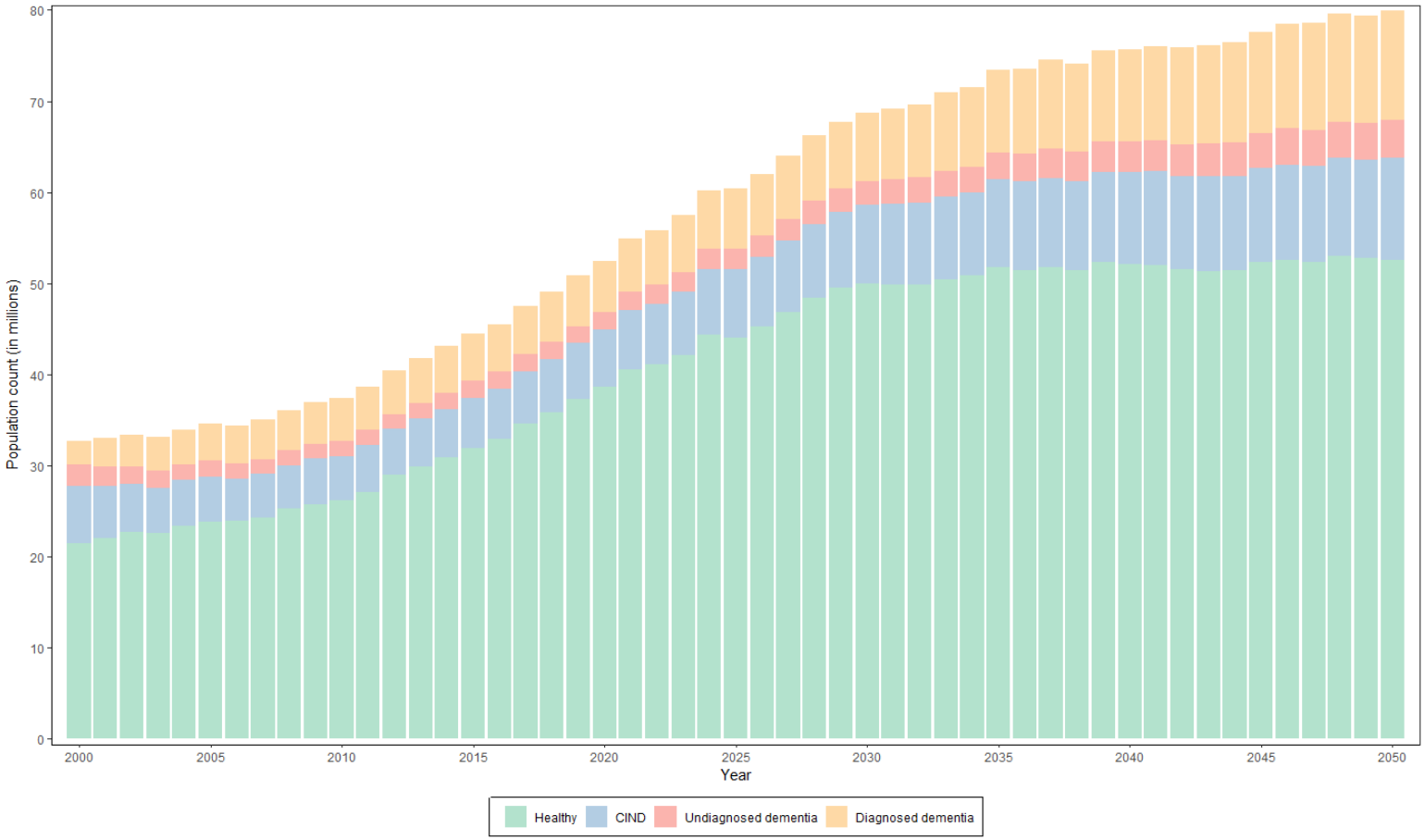


Figure 2.3 Forecasted population counts by cognitive status among adults aged 65 years or older, 2000-2050



Notes. CIND, cognitive impairment without dementia.

Figure 2.4 Forecasted prevalence of dementia among adults aged 65 years or older without (A) and with (B) age standardization to the population of 2000, 2000-2050

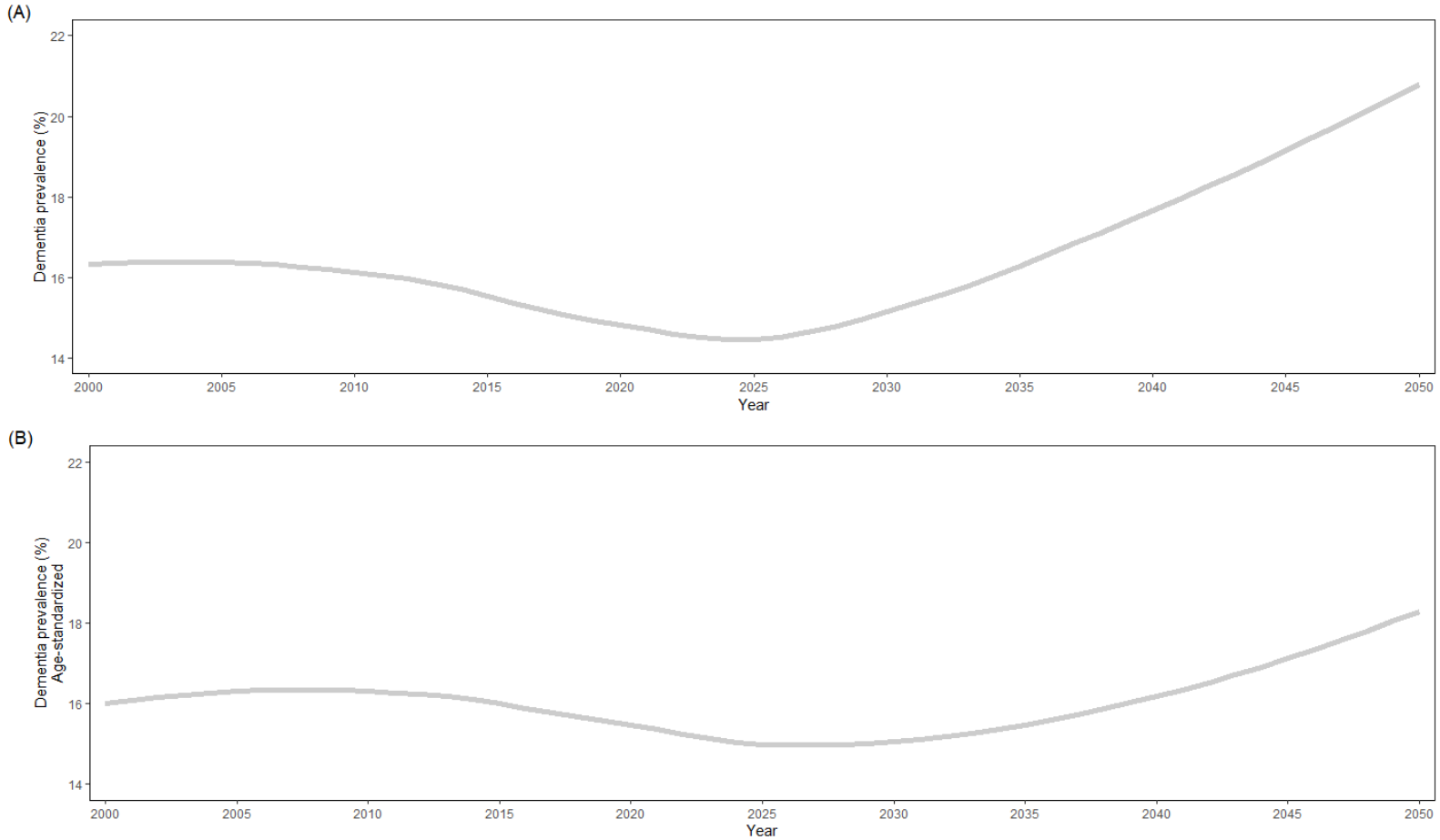
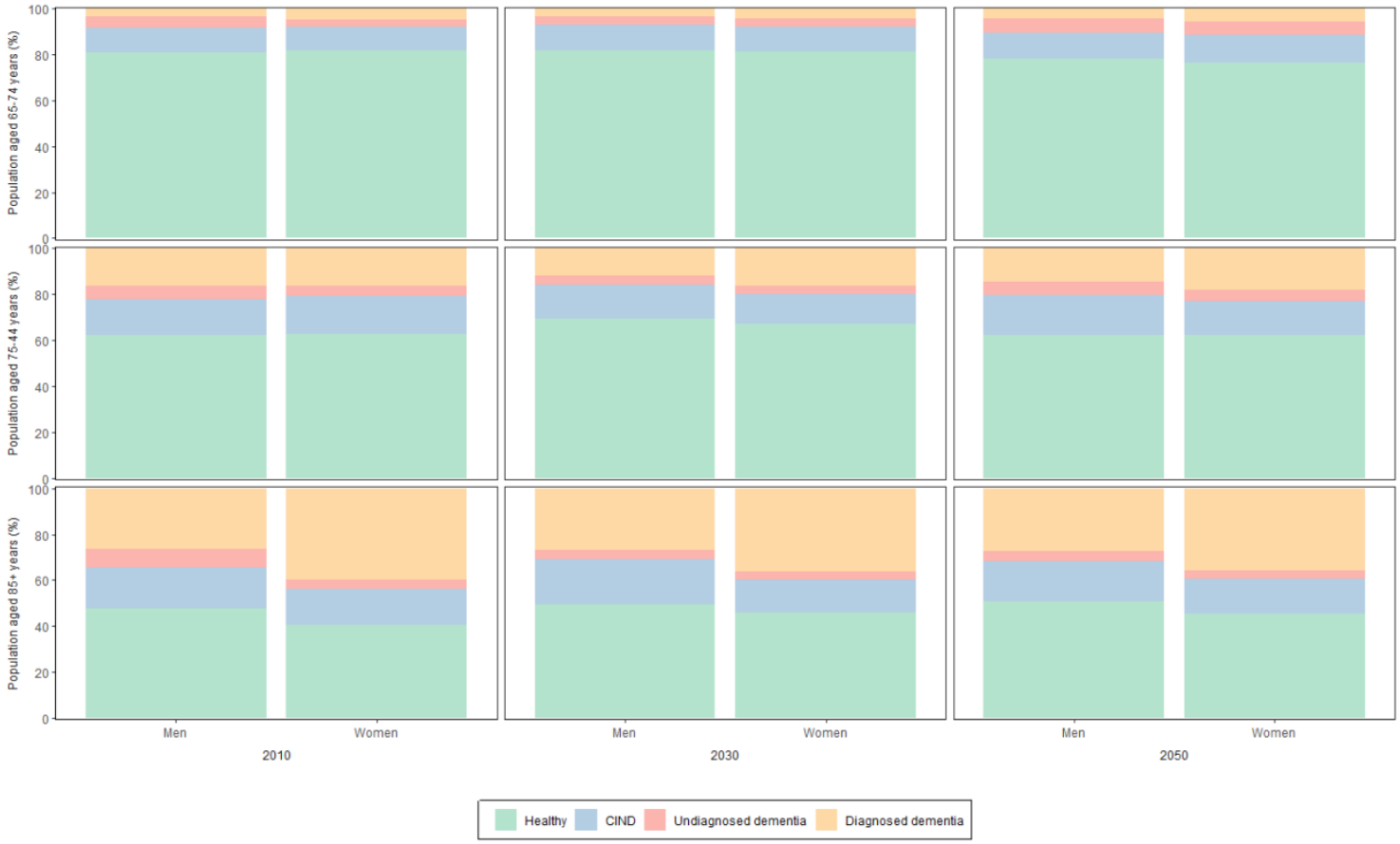
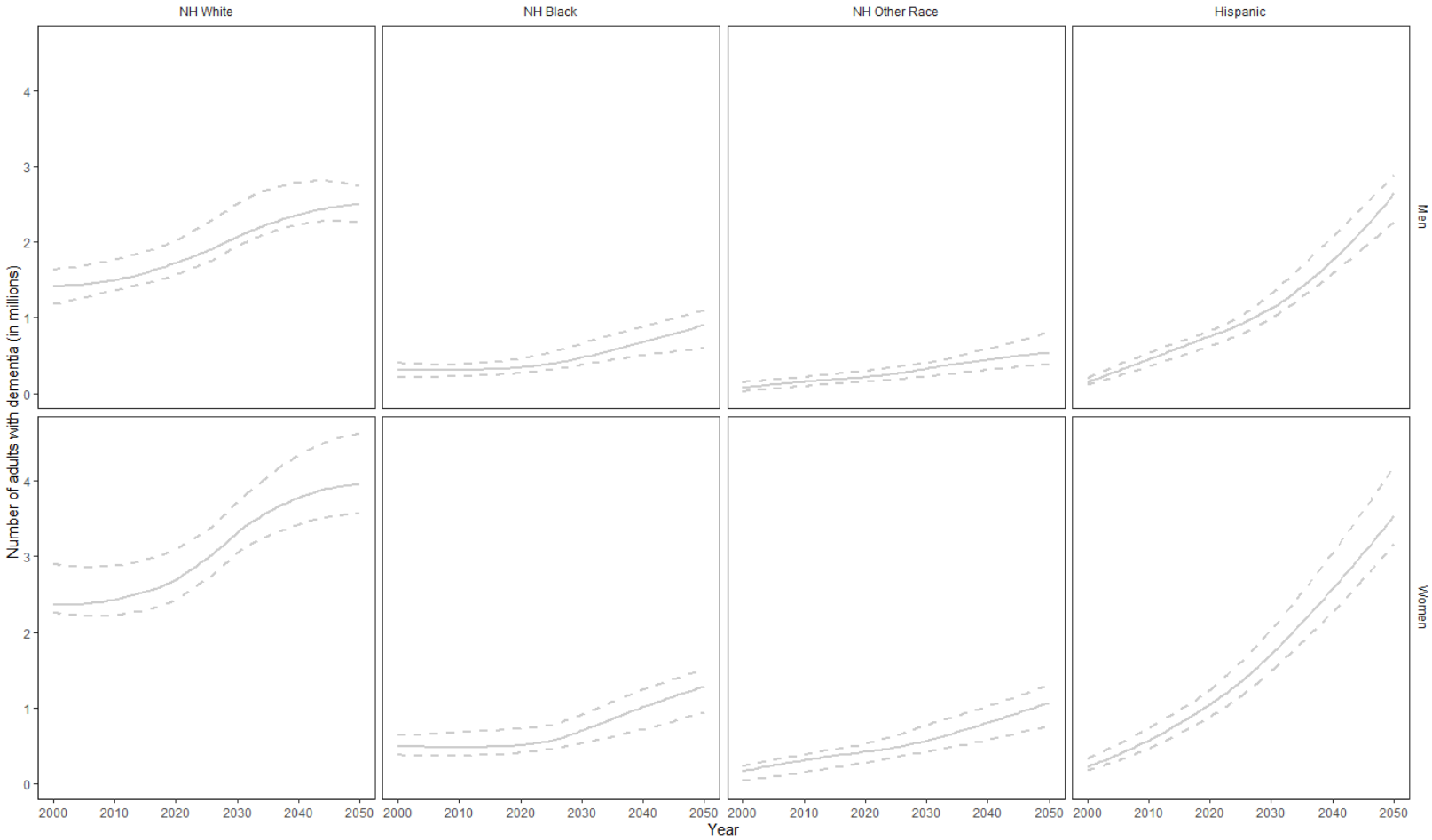


Figure 2.5 Proportion of adults without cognitive impairment or with CIND, undiagnosed dementia, or diagnosed dementia by sex and age group, 2010, 2030, 2050



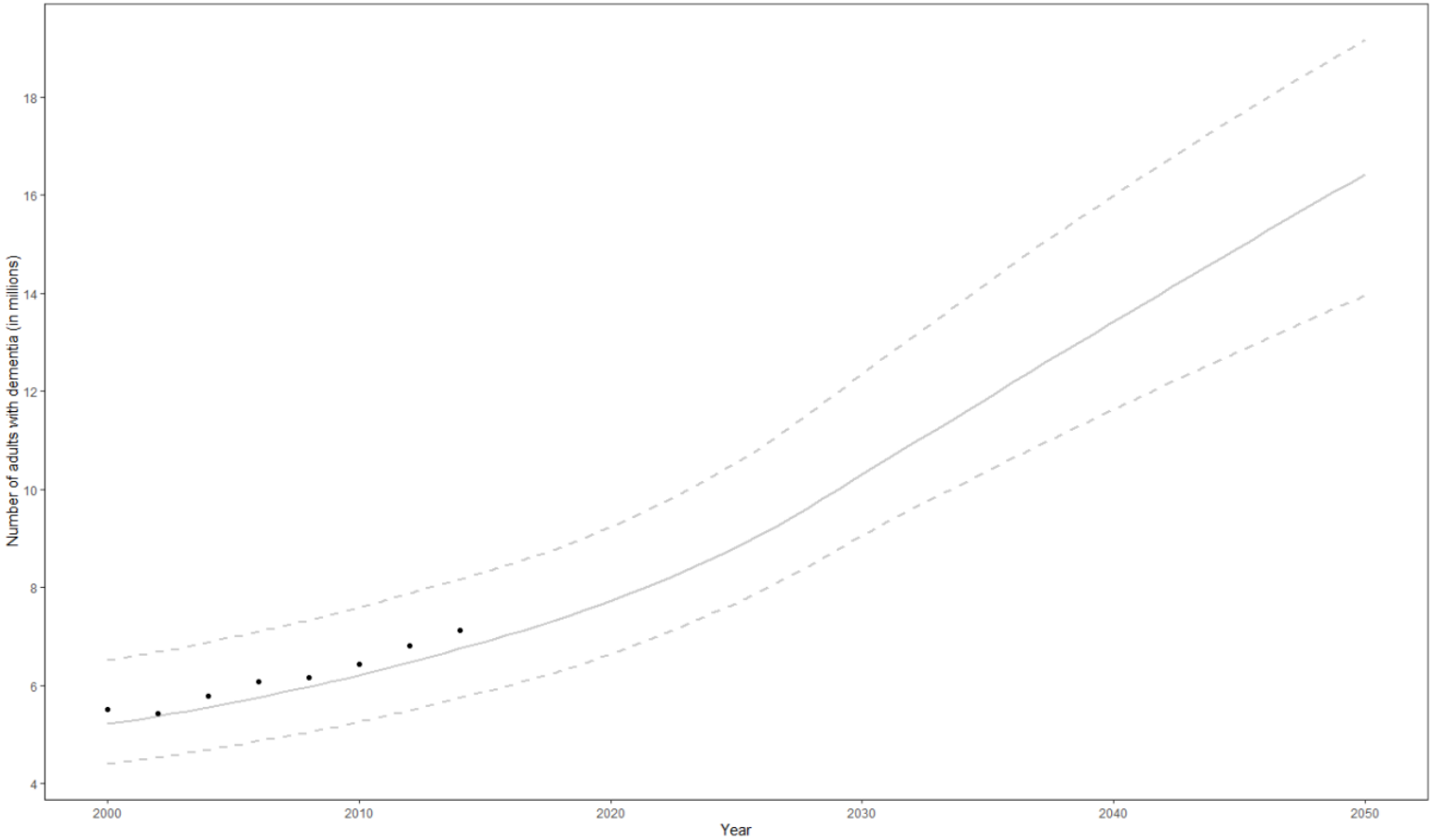
Notes. CIND, cognitive impairment without dementia.

Figure 2.6 Forecasted number of adults aged 65 years or older with dementia by sex and race/ethnicity group, 2000-2050



Notes. NH, non-Hispanic. Dashed lines correspond to the range from 500 simulations.

Figure 2.7 Actual and forecasted dementia, 2000-2050



Notes. Solid points indicate values obtained from HRS weighted to the Census. Dashed lines correspond to the range from 500 simulations.

Figure 2.8 Comparison of absolute (A) and logged (B) mean mortality rates obtained from the Dementia Population Cost Model with Human Mortality Database for the period 2000 to 2010

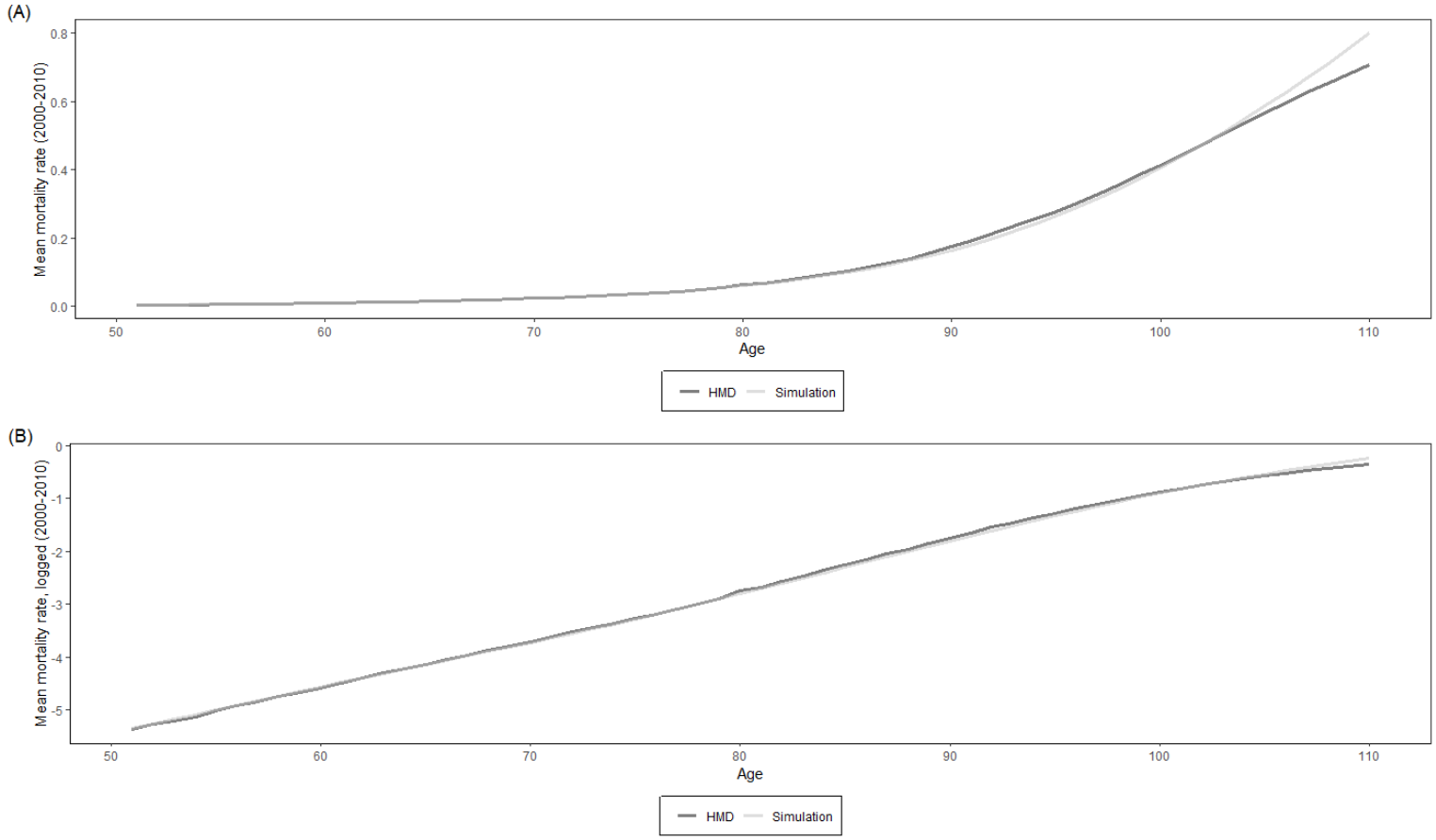


Figure 2.9 Comparison of population size obtained from the Dementia Population Cost Model with Census Projections

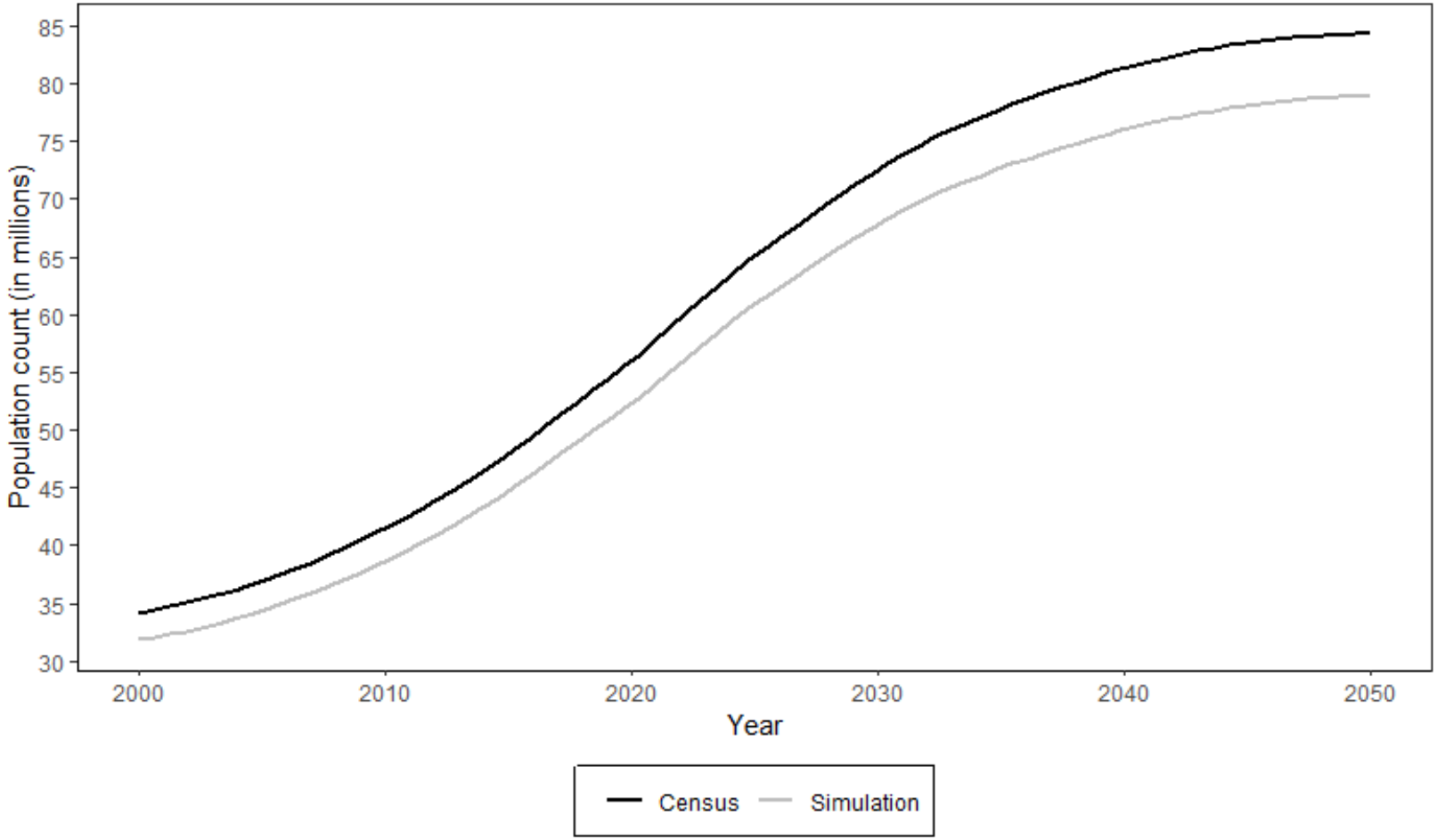
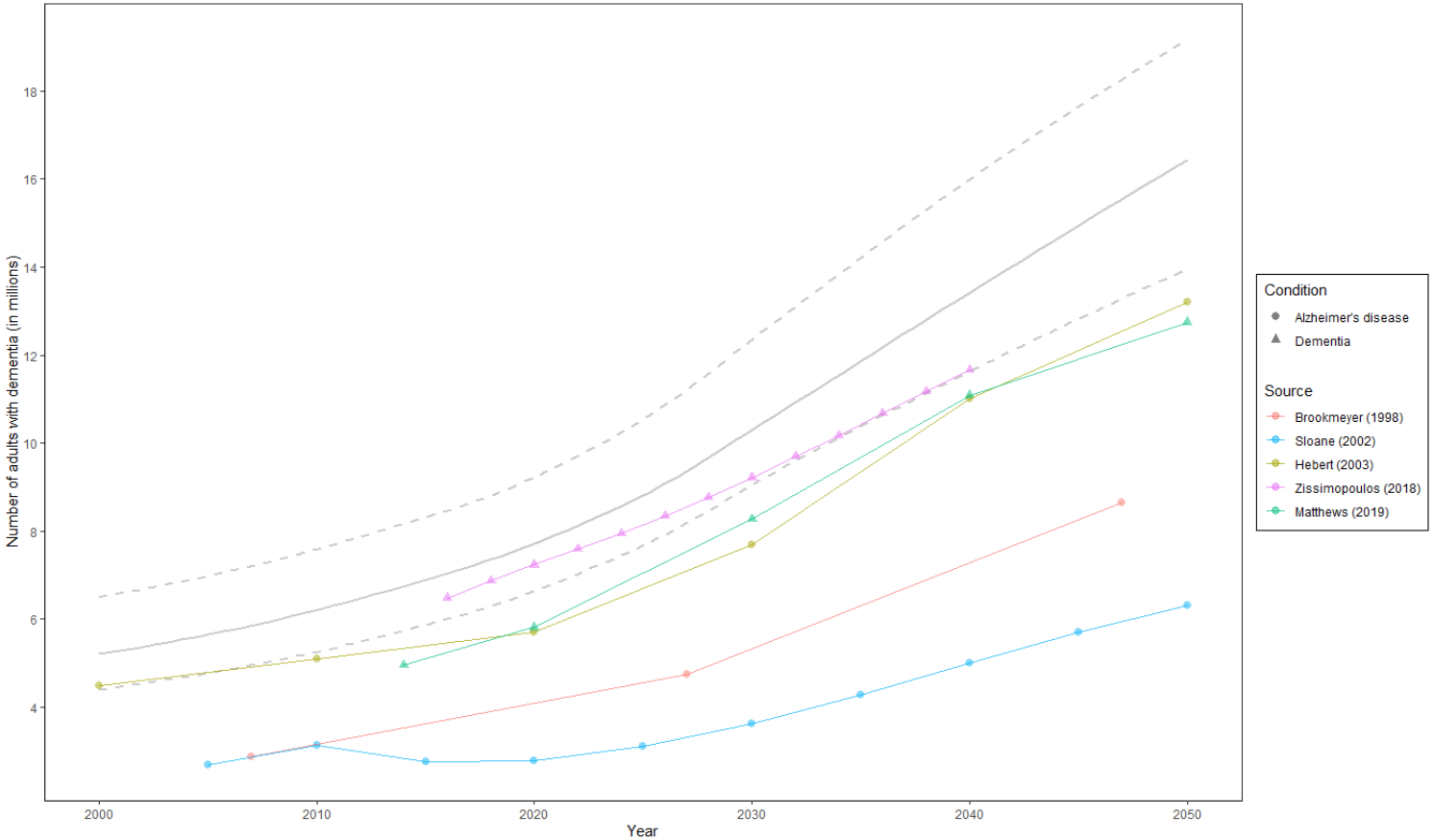


Figure 2.10 Comparison of forecasted dementia prevalence obtained from the Dementia Population Cost Model with previous studies



Notes. Results from the Dementia Population Cost Model are shown in gray. Dashed lines correspond to the range from 500 simulations.

2.7 Appendix

2.7.1 Tables

Table A.2.1 ICD-9 codes for dementia diagnosis in HRS-linked Medicare and Medicaid claims

ICD-9 Code	Description
331.0	Alzheimer's disease
331.1	Pick's disease
331.2	Senile degeneration of the brain
331.7	Cerebral degeneration in diseases classified elsewhere
290.0	Senile dementia, uncomplicated
290.1	Presenile dementia (brain syndrome with presenile dementia)
290.10	Presenile dementia, uncomplicated
290.11	Presenile dementia, with delirium
290.12	Presenile dementia, with delusional features
290.13	Presenile dementia, with depressive features
290.20	Senile dementia, with delusional features
290.21	Senile dementia, with depressive features
290.3	Senile dementia, with delirium
290.40	Arteriosclerotic dementia, uncomplicated
290.41	Arteriosclerotic dementia, with delirium
290.42	Arteriosclerotic dementia, with delusional features
290.43	Arteriosclerotic dementia, with depressive features
294.0	Amnestic syndrome (Korsakoff's psychosis or syndrome, nonalcoholic)
294.1	Dementia in conditions classified elsewhere
294.8	Other specified organ brain syndrome (chronic)
797	Senility without mention of psychosis

Table A.2.2 Model forms for transition models

Regrassand	Model form	Regressors
BMI	Multinomial logistic	age, sex, race, ethnicity, respondent's education, smoking status, BMI*
Diabetes	Binary logistic	age, sex, race, ethnicity, respondent's education, smoking status*, BMI*, diabetes*
Hypertension	Binary logistic	age, sex, race, ethnicity, respondent's education, smoking status*, BMI*, diabetes*, hypertension*
Heart disease	Binary logistic	age, sex, race, ethnicity, respondent's education, smoking status*, BMI*, diabetes*, hypertension*, heart disease*
Stroke	Binary logistic	age, sex, race, ethnicity, respondent's education, smoking status*, BMI*, diabetes*, hypertension*, heart disease*, stroke*
Cognitive status	Multinomial logistic	age, sex, race, ethnicity, respondent's education, smoking status*, BMI*, diabetes*, hypertension*, heart disease*, stroke*, cognitive status*

Notes. The transition probabilities were estimated using the state at time $t + 1$ as the dependent variable. The state at time t was used as a regressor.

* Indicates all available lagged terms were used.

Table A.2.3 Characteristics of the replenishing cohorts, 2010, 2020, 2030, 2040

Characteristic	2010	2020	2030	2040
N	4498619	4099623	4094877	4761958
Age				
51	4498619	4099623	4094877	4761958
Sex				
Male	2209609	2024796	2026072	2391954
Female	2289010	2074827	2068805	2370004
Race/ethnicity				
Non-Hispanic white	3162678	2556933	2301919	2534602
Non-Hispanic black	542761	494266	524249	663596
Non-Hispanic other	281328	348396	418336	553189
Hispanic	511852	700028	850373	1010571
Education, %				
< HS or GED	23.1	18.9	16.1	14.3
HS	29.3	27.8	23.3	21.5
Some college	27.1	28.2	31.0	30.8
College and above	20.5	25.1	29.6	33.4
Smoking status, %				
Never	42.1	49.6	53.8	57.2
Former	29.9	23.7	22.3	20.4
Active	28.0	26.7	23.9	22.4
BMI, %				
Underweight	1.4	0.8	0.6	0.6
Healthy weight	36.8	34.9	33.1	30.2
Overweight or obese	61.8	64.3	66.3	69.2
History of diabetes, %				
No	86.8	85.3	84.0	83.1
Yes	13.2	14.7	16.0	16.9
History of hypertension, %				
No	67.4	64.3	60.7	58.8
Yes	32.6	35.7	39.3	41.2
History of heart disease, %				
No	89.5	88.9	88.4	87.8
Yes	10.5	11.1	11.6	12.2
History of stroke, %				
No	96.7	95.8	94.9	94.1
Yes	3.3	4.2	5.1	5.9
State space, %				
No cognitive impairment	100	100	100	100
CIND	0	0	0	0
Undiagnosed dementia	0	0	0	0
Diagnosed dementia	0	0	0	0

Notes. BMI, body mass index; CIND, cognitive impairment without dementia; GED, general education development test; HS, high school.

Table A.2.4 Summary of assumptions for the Dementia Population Cost Model

Base Population
The sample-weighted HRS is nationally representative of the community-dwelling and nursing home populations in the US
State Space
The state space is comprehensive and each state corresponds to a weighted average across the spectrum of cognitive status for that given state
The state space accurately reflects the progression of cognitive impairment
Transition Models
Transitions between cognitive states and mortality as functions of covariates are relatively constant over time
The modeled relationships shown in Table A.2.2 are strictly additive
Transition probabilities for a given risk factor map to a weighted average across the spectrum of severity for that given risk factor in the population
The interpolation of the two-year transition probabilities into annual transition probabilities captures movements across states within one-year intervals
The order with which transition probabilities are estimated reflect the underlying pathology contributing to the progression of cognitive impairment
Immigrants to the US face the same risk set as native-born adults
One transition is made per simulation cycle
There are no transitions between education states after age 51
There are no transitions between smoking states after age 51
Diabetes and hypertension are absorbing states
Replenishing Cohorts
The number of 51-year-olds by sex and race/ethnicity entering the simulation in each cycle for the years 2000-2050 is in line with counts and forecasts provided by the US Census Bureau
The joint probability distribution of demographic characteristics, risk factors, and cognitive status is constant over time
The distribution of educational attainment, obesity, and smoking status by sex and race/ethnicity changes in accordance with trends observed in the BRFSS and, in turn, trends in medical risk factors, cognitive status, and mortality are accurately reflected
The distribution of educational attainment, obesity, and smoking status by sex and race/ethnicity changes in accordance with trends observed in the BRFSS and, in turn, trends in medical risk factors, cognitive status, and mortality are accurately reflected
All entering 51-year-olds enter the model without cognitive impairment

Table A.2.5 Summary of existing dementia projections and forecasts

First Author	Year	Condition	Period	Method	Data Sources	Model severity
Brookmeyer	1998	Alzheimer's disease	1997-2047	Macrosimulation based on incidence model	US Census Bureau population projections; Framingham Heart Study; East Boston Study; Rochester Study; Baltimore Longitudinal Study of Aging	No
Sloane	2002	Alzheimer's disease	2000-2050	Macrosimulation based on incidence model	US Census Bureau population projections; estimates of AD provided by the U.S. General Accounting Office	Yes
Hebert	2003	Alzheimer's disease	2000-2050	Macrosimulation based on incidence model	US Census Bureau population projections; Chicago Population Study	No
Matthews	2019	Dementia	2015-2060	Extrapolation based on prevalence	US Census Bureau population projections; Prevalence of dementia among Medicare Fee-for-service beneficiaries	No
Hurd	2013	Dementia	2010-2040	Extrapolation based on predicted probability	US Census Bureau population projections; Health and Retirement Study	No
Zissimopoulos	2018	Dementia	2016-2040	Microsimulation	US Census Bureau population projections; Health and Retirement Study; National Health Interview Survey	No

Notes. I searched the PubMed database in April 2019 to review existing evidence for studies examining future dementia trends in the US. The following search query was used: "Dementia"[Mesh] AND "United States"[Mesh] AND ("Simulation" OR "Forecast" or "Projection"). I manually removed research articles that were not relevant and performed additional searches using reference lists for the selected papers.

2.7.2 Figures

Figure A.2.1 Comparison of weighted HRS base population with Census using population counts (A) and ratios (B) by age group, 2000

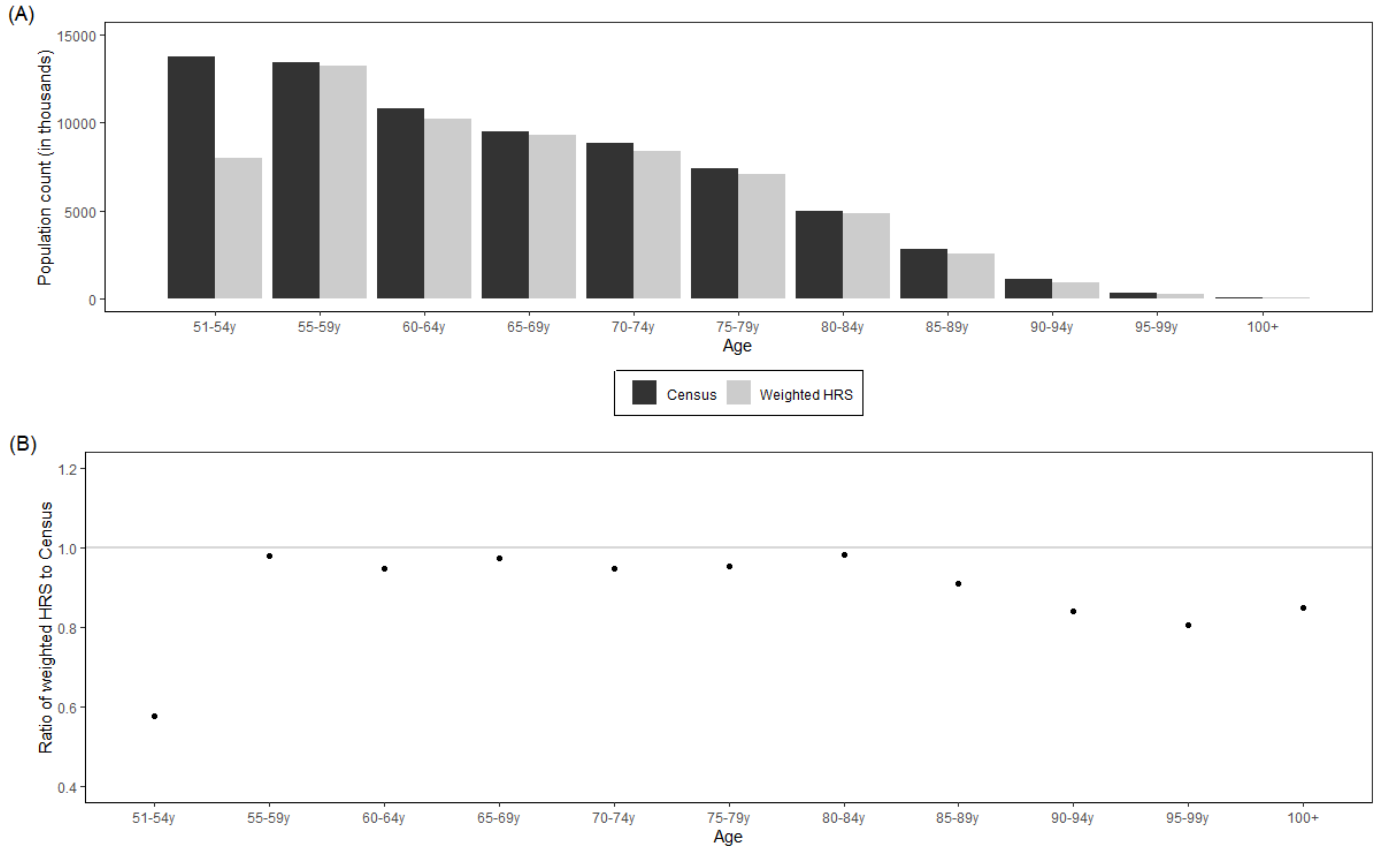
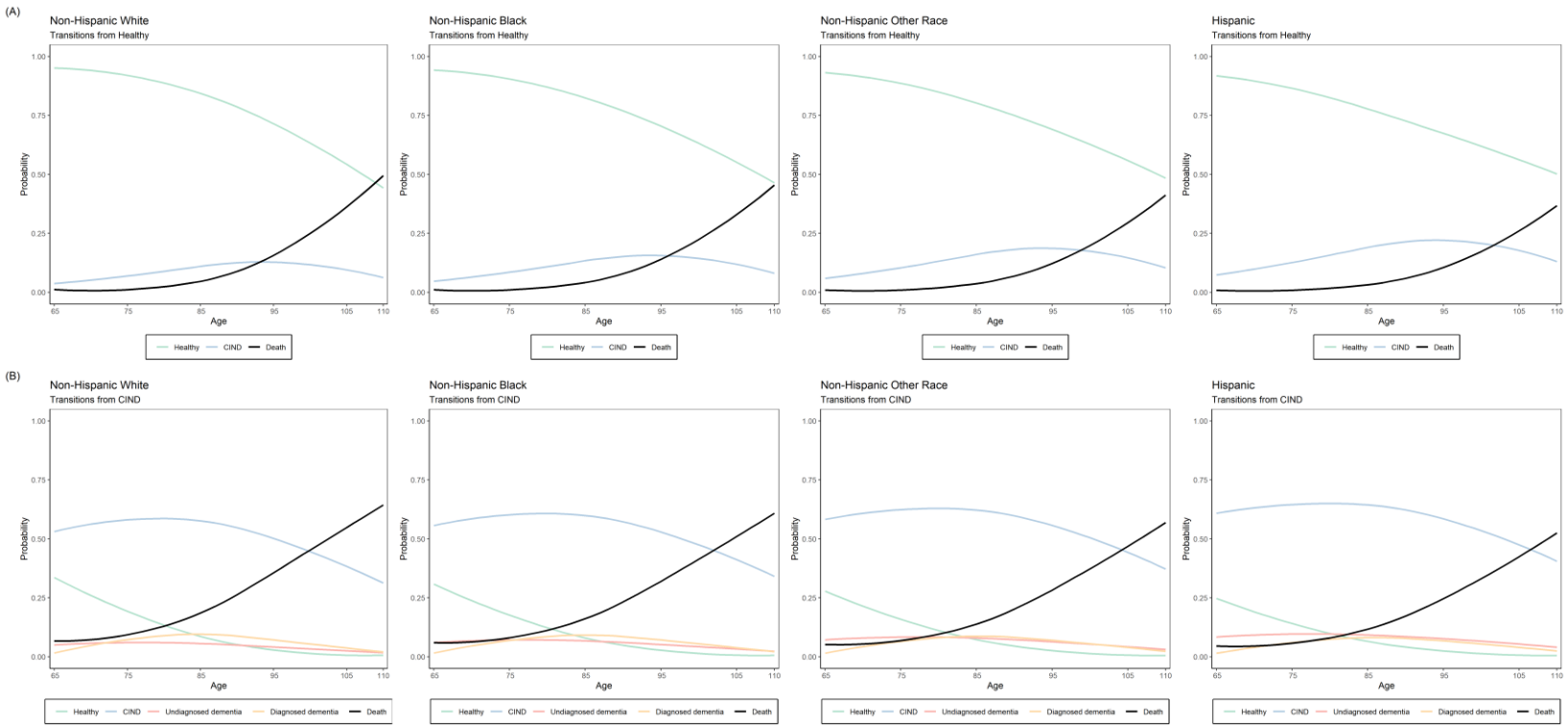


Figure A.2.2 Sample transition probabilities by race/ethnicity from initial states of healthy (A), CIND (B), Undiagnosed dementia (C), and Diagnosed dementia (D)



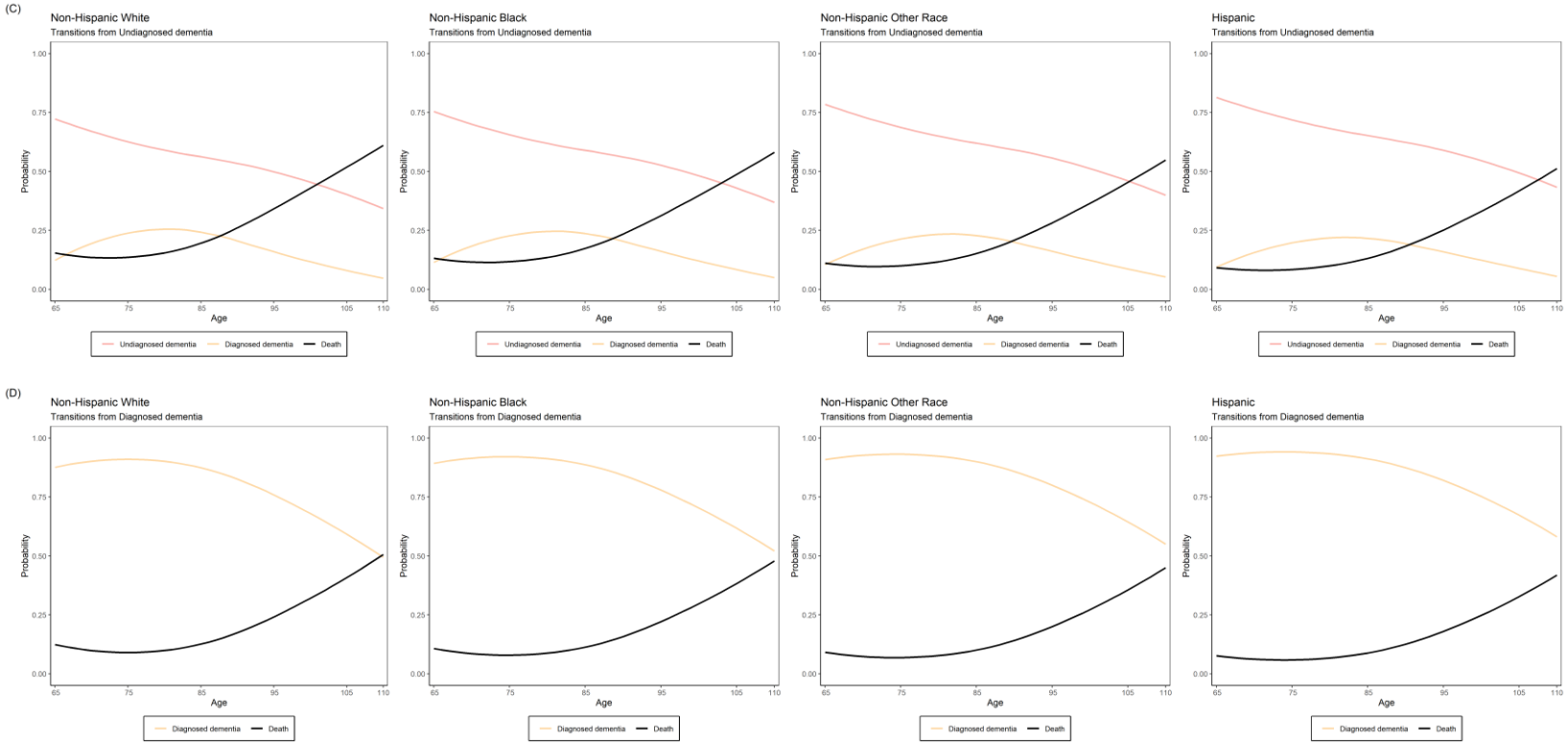
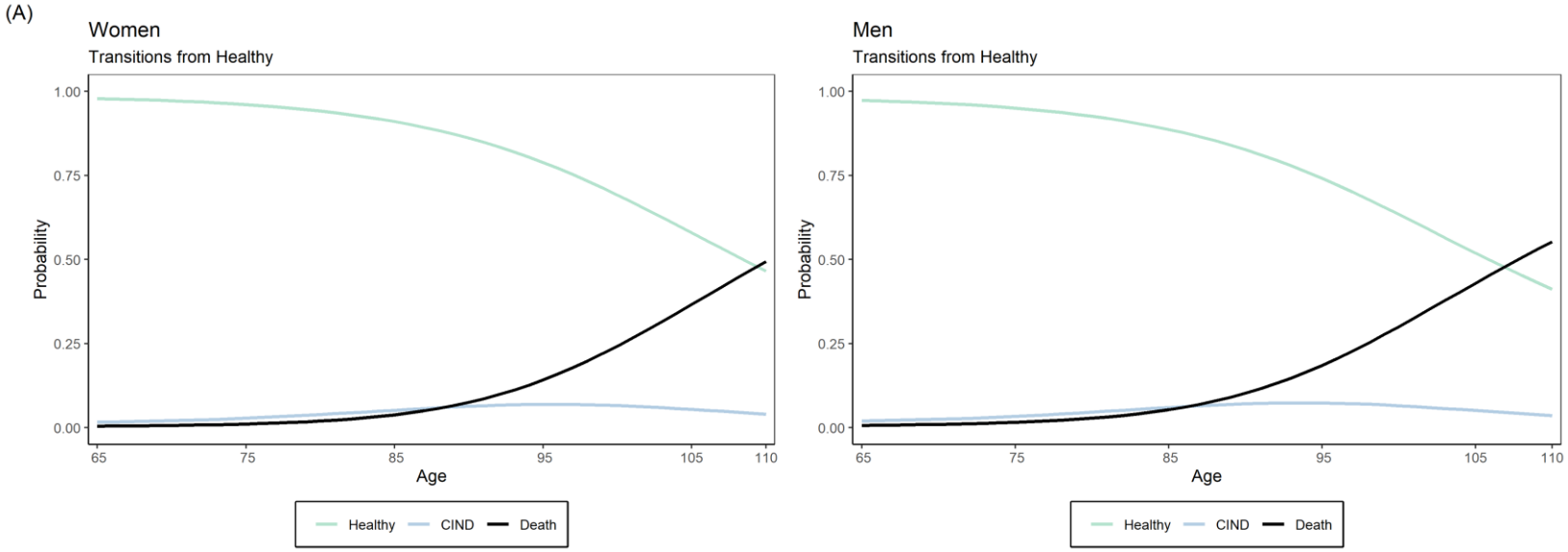
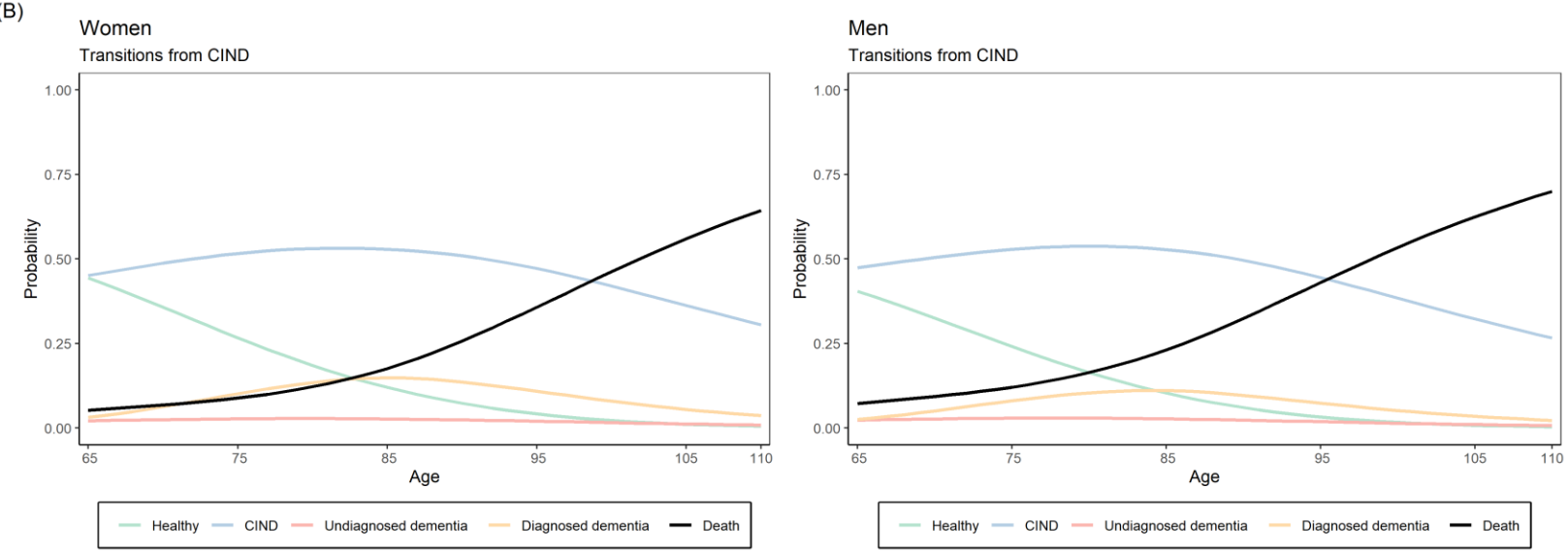

(C)
Non-Hispanic White
Transitions from Undiagnosed dementia
Non-Hispanic Black
Transitions from Undiagnosed dementia
Non-Hispanic Other Race
Transitions from Undiagnosed dementia
Hispanic
Transitions from Undiagnosed dementia
Probability
Age
Undiagnosed dementia
Diagnosed dementia
Death
(D)
Non-Hispanic White
Transitions from Diagnosed dementia
Non-Hispanic Black
Transitions from Diagnosed dementia
Non-Hispanic Other Race
Transitions from Diagnosed dementia
Hispanic
Transitions from Diagnosed dementia
Probability
Age
Diagnosed dementia
Death

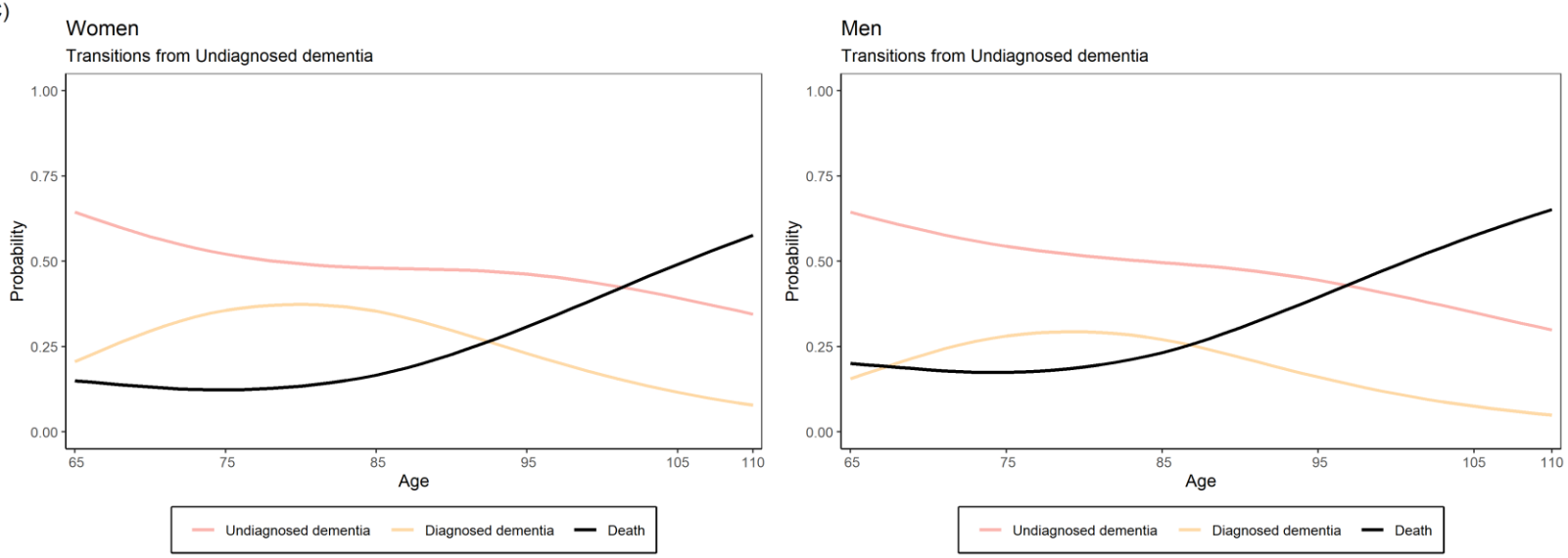
Figure A.2.3 Sample transition probabilities by sex for Non-Hispanic white adults with a college level of educational attainment from initial states of healthy (A), CIND (B), Undiagnosed dementia (C), and Diagnosed dementia (D)



(B)


Women
Transitions from CIND
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Healthy
CIND
Undiagnosed dementia
Diagnosed dementia
Death
Men
Transitions from CIND
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Healthy
CIND
Undiagnosed dementia
Diagnosed dementia
Death

(C)


Women
Transitions from Undiagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Undiagnosed dementia
Diagnosed dementia
Death
Men
Transitions from Undiagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Undiagnosed dementia
Diagnosed dementia
Death

(D)

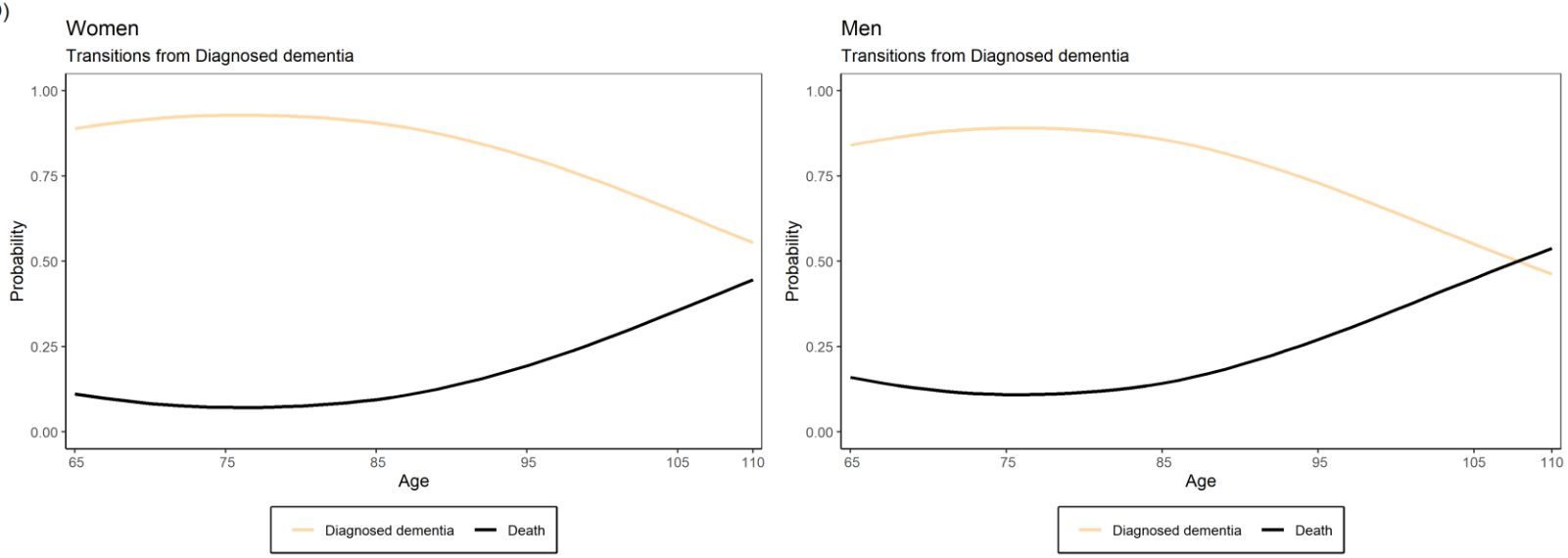
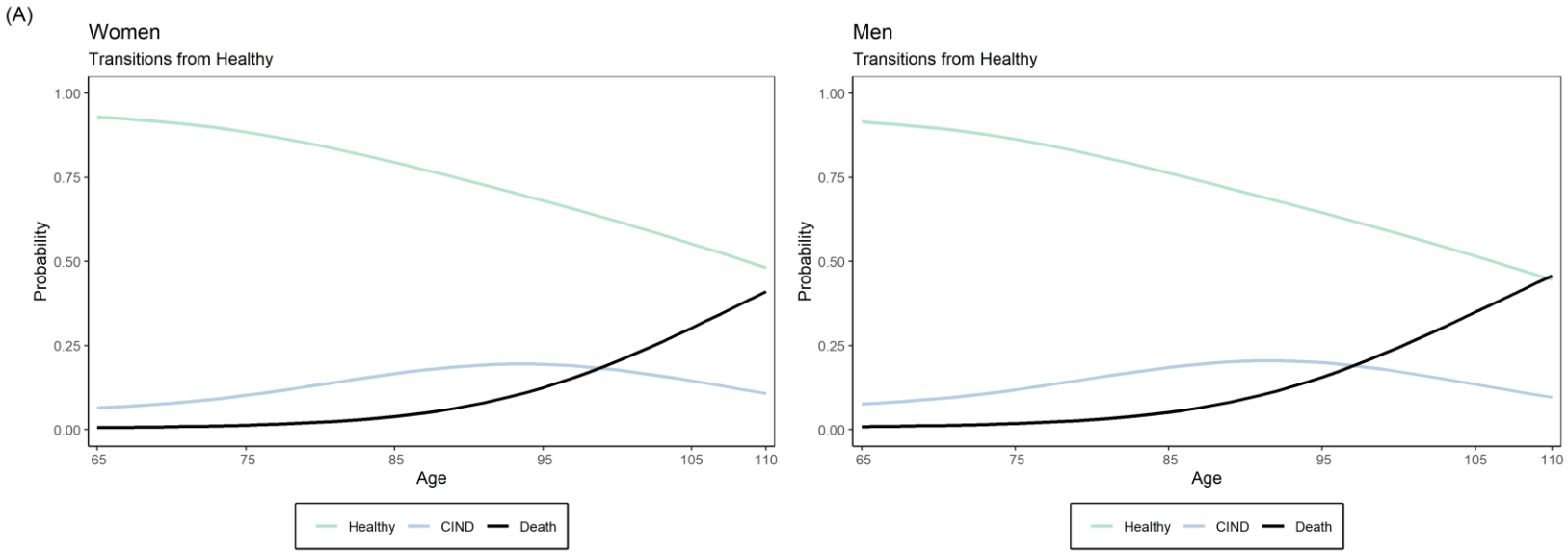
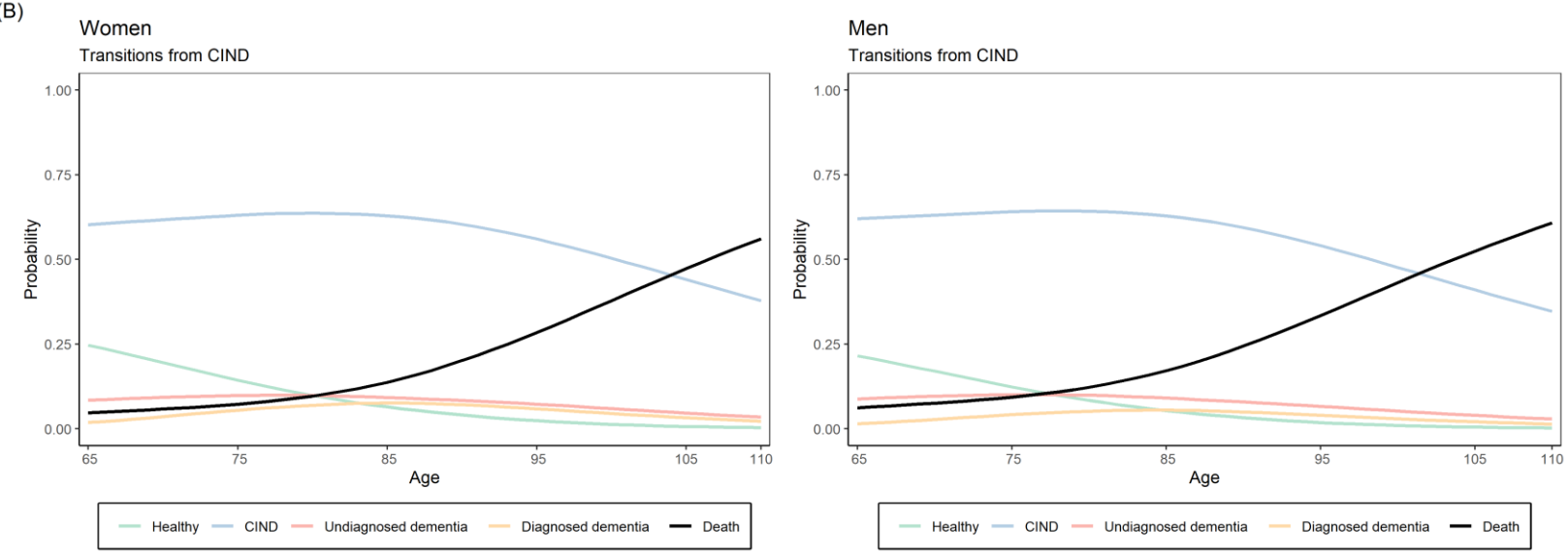

Women
Transitions from Diagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Diagnosed dementia
Death
Men
Transitions from Diagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Diagnosed dementia
Death

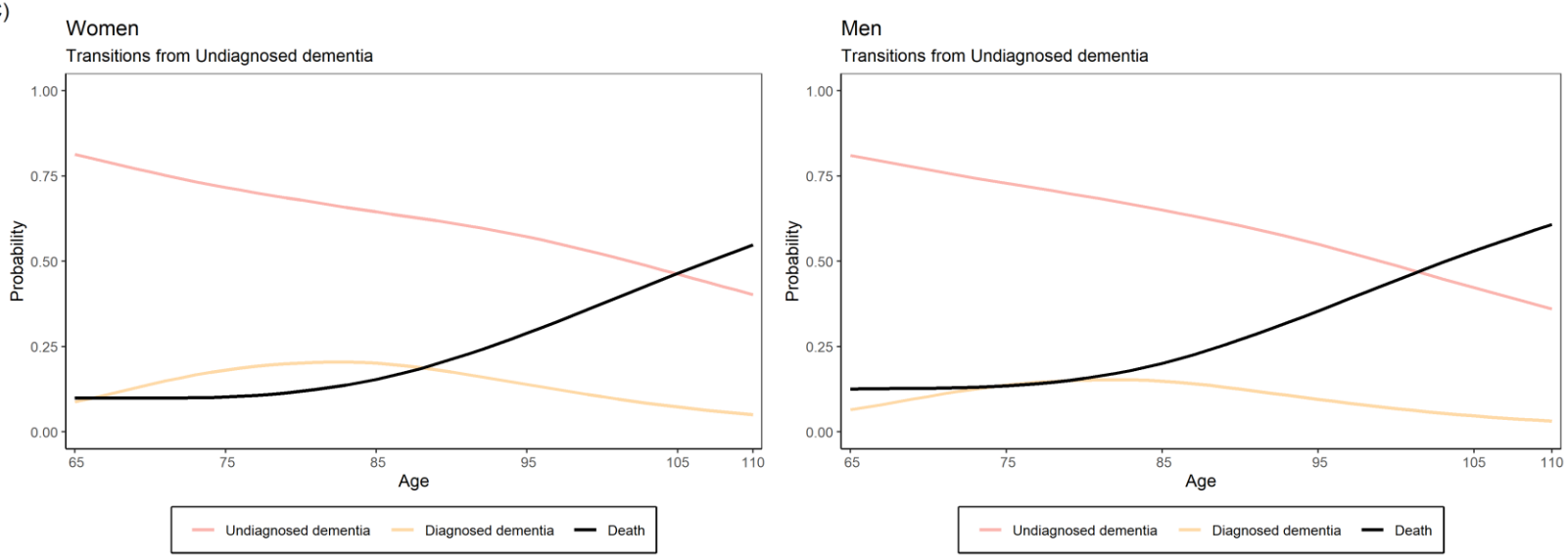
Figure A.2.4 Sample transition probabilities by sex for Non-Hispanic white adults with less than a high school degree or GED from initial states of healthy (A), CIND (B), Undiagnosed dementia (C), and Diagnosed dementia (D)



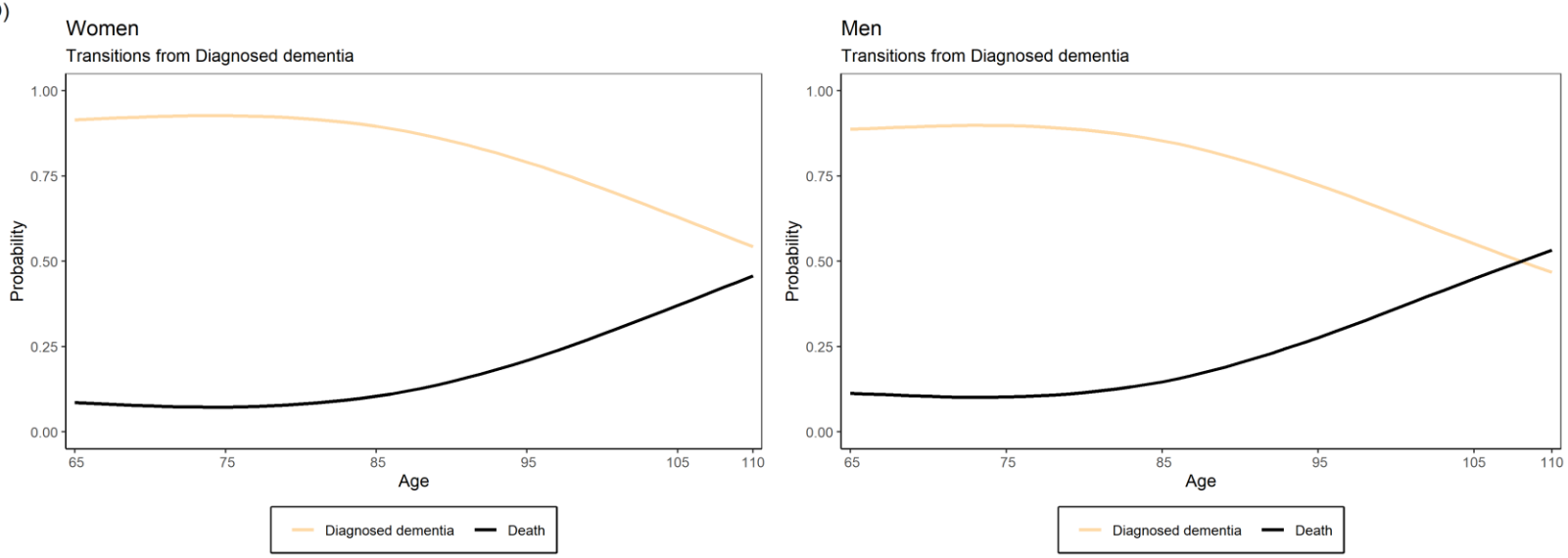
(B)


Women
Transitions from CIND
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Healthy
CIND
Undiagnosed dementia
Diagnosed dementia
Death
Men
Transitions from CIND
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Healthy
CIND
Undiagnosed dementia
Diagnosed dementia
Death

(C)


Women
Transitions from Undiagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Undiagnosed dementia
Diagnosed dementia
Death
Men
Transitions from Undiagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Undiagnosed dementia
Diagnosed dementia
Death

(D)


Women
Transitions from Diagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Diagnosed dementia
Death
Men
Transitions from Diagnosed dementia
Probability
1.00
0.75
0.50
0.25
0.00
65
75
85
95
105
110
Age
Diagnosed dementia
Death

CHAPTER 3: Forecasting the Cost of Dementia in the United States

3.1 Introduction

Dementia imposes a substantial burden on society in the form of direct and indirect medical costs as well as intangible costs in the form of reduced quality of life for patients, their families, and caregivers. The estimated economic burden of dementia in the United States in 2019 was $290 billion, of which $195 billion was covered by Medicare and Medicaid (Alzheimer's Association 2019). These costs are expected to rise in response to projected increases in the number of adults with dementia and will be influenced by changes in the demographics and comorbidities of the population with dementia, health care utilization and delivery patterns, and evolution of payer mix.

The economic burden of dementia can be measured through direct, indirect, and in some cases, intangible costs (Byford, Torgerson and Raftery 2000). Direct costs are those related to institutional care and include medical services and treatment, medications, formal care, and institutional care. Indirect costs associated with dementia, such as unpaid (i.e., informal) caregiving by a family member or friend as well as wages lost by the patient or informal caregiver, are considered a key driver of the burden of dementia due to the increased care requirements for dementia patients; nearly half of all informal caregivers in the United States provided care for a patient with dementia (Friedman et al. 2015). Intangible costs tend to be more qualitative in nature and include, for example, suffering and quality of life of the patient and caregiver; intangible costs are often difficult to quantify and are thus, typically excluded from cost estimates.

Several studies over the past two decades have characterized and projected the monetary costs of dementia in the US with discrepant results (Bharmal et al. 2012;

Bynum et al. 2004; Deb et al. 2017; Frytak et al. 2008; Goldfeld et al. 2011; Hurd et al. 2013; Joyce et al. 2007; Kelley et al. 2015; Kuo et al. 2008; Suehs et al. 2014; Taylor Jr and Sloan 2000; Weiner et al. 1998; White et al. 2019; Yang and Levey 2015; Zhu et al. 2015a; Zhu et al. 2015b; Zhu et al. 2006; Zhu et al. 2008; Zissimopoulos, Crimmins and St Clair 2014). The few cross-sectional studies reporting Medicare expenditures after the first year of dementia diagnosis even vary, ranging from $2,752 (Hurd et al. 2013) to $10,598 (Bynum et al. 2004) per person.

The notable discrepancy in reported costs likely stems from variation in study design as well as what these costs map to. Most studies vary with respect to case definition (e.g., Alzheimer's disease-dementia versus all-cause dementia) and case ascertainment (e.g., clinical diagnosis, survey-based), the time of diagnosis and the stage at which costs were captured (e.g., prior to onset, at onset, post-onset), sample ascertainment (e.g., survey based, administrative health records), source of costs (e.g., out-of-pocket, Medicare, Medicaid, informal care), as well as cost estimation approaches (e.g., matched study, counterfactual simulation) and types of costs (e.g., lifetime, prevalent, incremental).

Typically, but not consistently, researchers assess the incremental (i.e., additional) cost of health care attributable to dementia by comparing costs among patients with dementia and costs among patients who are dementia-free; the resulting difference being the incremental cost. Other studies estimate incremental costs by comparing costs among dementia patients to what their costs would have been in a potential outcomes framework in which that same individual did not have dementia (e.g., by way of counterfactual simulation). In the context in which researchers compare cases (i.e., patients with dementia) to controls (i.e., dementia-free patients), there is a great deal of variation with

respect to the criteria used for comparisons. Some studies compare cases and controls who are similar along a set of baseline demographic characteristics and comorbid conditions whereas other studies compare average costs among cases to average costs among controls—making to adjustments for such factors as education or coexisting medical conditions. Studies using the same data source but different study designs have also reported discrepant results.

Hurd and colleagues (Hurd et al. 2013) estimated the total cost of prevalent dementia among adults older than 70 years of age in 2010 using data from the Health and Retirement Study. To isolate the per person costs attributable to dementia, Hurd et al. estimated the probability that an individual had dementia using a regression-based probit model. Then, costs were compared among individuals with a predicted dementia probability of zero (dementia-free controls) and those with a predicted probability of one (cases). After adjustment for demographic characteristics and comorbid conditions, the annual per-person cost for care purchased in the market attributable to dementia was $28,501 (95% confidence interval: $20,881 to $36,122), of which approximately $2,700 was borne by Medicare. Cost of care purchased in the market was estimated using self-reported out-of-pocket spending, Medicare spending, nursing home spending, and spending on in-home care, amounting to an estimated $109 billion or between 51% and 69% of total estimated costs. Two different approaches were used to estimate indirect costs which were reflected in the range of estimates: the replacement cost approach, which yielded the higher estimate and was responsible for 49% of total costs, and the foregone wages approach which was responsible for 31% of total costs. The replacement cost approach estimates the monetary value of care by using the cost of an equivalent

service purchased in the market through a home health agency (National Research Council 2005). The forgone wage approach bases the valuation on the labor-market income forgone because of time spent on caregiving (Glied 1996). The use of either approach illustrates the large contribution of informal care to the overall cost of dementia.

Using the same data source, period (i.e., Hurd focused used year 2010; the study by Kelley spanned 2005 to 2010), age criterion, and classification scheme for dementia as Hurd and colleagues (Hurd et al. 2013), Kelley et. al. (Kelley et al. 2015) estimated the five-year cost of dementia prior to death specifically among Medicare fee-for-service (FFS) beneficiaries. Among the 555 decedents with dementia, the mean per-person Medicare costs were $86,430 over the five year period prior to death. This estimate is notably higher than Medicare-specific cost estimates reported by Hurd. However, overall, the cost estimates reported by Kelley et. al. were noticeably lower than those reported by Hurd et. al. In addition, in the Kelley article, average Medicare expenditures were found to be comparable among the dementia-free and dementia subgroups. This discrepancy is most likely attributable to the fact that the study by Kelley compared costs in the last five years of life; thus, because all included participants in their sample were within five years of death, they were all likely to be in poor health irrespective of dementia status.

Another study (Zissimopoulos et al. 2014), also using data from the Health and Retirement study, but a different classification scheme for dementia and obtaining Medicare cost estimates from the 2002-2004 Medical Expenditure Panel Survey and Medicare Current Beneficiary Survey—specifically for Alzheimer's disease—reported a per-person annual Medicare cost of $17,444 in 2010 compared to $10,904 among individuals not classified with dementia.

As illustrated, three contemporary studies using the same data source but different study designs reported varying cost estimates of dementia which map to different cost estimates (e.g, prevalent costs versus costs in the last five years of life). Other recent studies estimated the lifetime costs for dementia patients using a counterfactual simulation approach. Jutkowitz and colleagues (Jutkowitz et al. 2017) developed a model to estimate the lifetime costs of dementia compared to dementia-free individuals by simulating disease progression for newly diagnosed dementia patients. The average net lifetime cost attributable to dementia for an 83-year-old with a projected five-year survival after diagnosis was $321,880 (in 2015 dollars); $184,500 more than the lifetime costs of formal and informal healthcare for non-dementia individuals. The key driver of this difference was the value of informal caregiving required for dementia patients. The authors also found the net cost for dementia patients to vary in the years following diagnosis, with annual costs in the first five years being driven by the value of informal care, and out-of-pocket expenses due to long-term care facility representing an increasing share of costs in the later five years. In addition, the cost of care varied by the number of functional limitations, which was also found in a recent study using data from the Health and Retirement Study (Coe, Skira and Larson 2018).

In addition to the higher costs associated with informal care, dementia patients compared to non-dementia patients utilize health care services at a higher rate. For example, dementia patients visit hospitals at twice the rate of non-dementia patients, are four times as likely to require skilled nursing facilities (SNFs), and are more likely to have a least one home health care visit per year (Alzheimer's Association 2019). This is due, in part, to the increased likelihood of a dementia patient having a comorbid

condition, such as diabetes or a stroke, which are prevalent among an estimated 37% and 22% of dementia patients, respectively (Alzheimer's Association 2019). Thus, future dementia costs will be influenced by changes in risk profiles and comorbidities, which have typically been left out of existing models projecting the economic burden of dementia.

Although several studies have examined costs associated with dementia, few have examined how these costs are met. Approximately one quarter of the lifetime costs of dementia are covered by individuals out-of-pocket and nearly 67% are covered by public programs, such as Medicare and Medicaid (Alzheimer's Association 2019). Medicare FFS is a public health insurance program for adults age 65 and older with three main parts: Parts A, B, and D (Centers for Medicare Medicaid Services 2018). Medicare Part A covers up to 90 days of inpatient hospital stays and up to 100 days of a SNF during the beneficiary's lifetime. Medicare Part B requires a monthly premium that provides medical insurance coverage for physician visits and outpatient hospital procedures. Medicare Part D is used to help pay for self-administered prescription drugs through prescription drug insurance premiums. Although Medicare coverage includes the diagnosis and partial treatment of dementia, it does not cover long term care which represents the largest expense for dementia patients (Alzheimer's Association 2019). Medicaid, another form of public health coverage, does cover long term care but is restricted to low income individuals (regardless of age) and those with disabilities (Feder, Komisar and Niefeld 2000; Ng, Harrington and Kitchener 2010). In 2018, Medicare and Medicaid covered an estimated $140 billion and $47 billion of expenses respectively for dementia patients. Approximately 1 in every 5 Medicare dollars was spent on dementia in

2018, and this is expected to rise to nearly 1 in every 3 Medicare dollars by 2050 (Alzheimer's Association 2019). To date, studies of dementia cost have typically derived cost estimates from Medicare FFS claims data applied to the prevalence of dementia. Few studies have integrated periodized costs borne by Medicare with out-of-pocket expenses. Even less is known about how cost structures may vary among adults who are clinically diagnosed with dementia compared to those who have a dementia-level of impairment but remain undiagnosed.

With rising health care costs and limited resources, it is necessary to understand the economic cost of dementia and its determinants to inform health policy and health care resource allocation. However, few studies have provided dementia cost projections that effectively synthesize dynamic projections of the prevalence, incidence, and mortality of dementia while accounting for demographic and epidemiological forces. This study has two objectives. First, we provide contemporary estimates of the incremental direct costs of undiagnosed dementia and diagnosed dementia. We focus on out-of-pocket expenditures and costs borne by Medicare rather than comprehensive care costs. Second, we combine these estimates with forecasted population counts for each level of severity reported in Chapter 2 to estimate the future direct costs of dementia in the US.

3.2 Materials and Methods

3.2.1 Data

The economic burden of dementia was based on estimates from multiple data sources.

3.2.2 Dementia Population Cost Model

Annual population counts for the years 2000 to 2050 of cognitive impairment by severity (undiagnosed dementia, diagnosed dementia) were obtained from the Dementia Population Cost Model. These prevalence estimates were disaggregated by age, sex (male, female), race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, or Hispanic), and educational attainment (less than high school or GED, high school, some college, college and above). The Dementia Population Cost Model also tracks in which year an individual enters a given state of impairment and for how long they remain in that state.

3.2.2.1 Health and Retirement Study and Linked Claims Data

We used data from the Health and Retirement Study (HRS) linked with respondent's Medicare claims for the years 1991 through 2012. We obtained Medicare expenditures for the 12 months prior and up to 60 months following a diagnosis of dementia as recorded on claims submitted for reimbursement. In addition, we obtained out-of-pockets expenditures through the HRS survey data. Study procedures were approved by the Institutional Review Board at the University of Pennsylvania, the HRS Restricted Data Applications Processing Center, and the Centers for Medicare and Medicaid Services (CMS) Privacy Board.

3.2.3 Sample ascertainment

The analytic sample was comprised of the subset of HRS respondents who consented to review of their medical records (~80%) and had fee-for-service (FFS) coverage at least 12 months prior to and one month after diagnosis. In total, our analytic sample was comprised of 22,339 HRS respondents who provided their Medicare

identification number and for whom we obtained validated enrollment information for the years 1991 to 2012. Among the 22,339 eligible respondents, 5,477 (24.5%) had a dementia diagnosis code; 4,335 (19.4%) also met the Medicare enrollment criteria.

3.2.4 Measures

We identified participants with dementia using ICD-9-CM diagnostic codes from Medicare claims. In our classification, we required that patients were enrolled in Medicare parts A and B FFS coverage for at least 12 months before and one month after receipt of one of the following diagnosis codes in an inpatient, skilled nursing facility, home health, hospital outpatient, or carrier claim: 331.x, 290.x, 294.x, or 797 (Centers for Medicare & Medicaid Services 2018; Lin et al. 2016).

Self-reported out-of-pocket expenditures from HRS survey respondents were combined with health service utilization and cost data obtained through linked CMS records. These expenditures included costs covered by Medicare (parts A and B), including inpatient, skilled nursing facility, institutional outpatient, carrier, home health, durable medical equipment, and hospice services. We excluded expenses for prescription drugs not administered by a physician. Expenditures were examined 12 months prior to and up to 60 months following a dementia diagnosis as ascertained through Medicare claims submitted for reimbursement. We carried forward costs in the fifth year for individuals with impairment who survived beyond five years. End of life costs were obtained through exit interviews conducted with a family member of the deceased HRS respondent.

3.2.5 Statistical Analysis

A key challenge to estimating the incremental cost of dementia is one of comparing potential outcomes. Ideally, we would compare the costs for dementia patients to what their costs would have been if they were dementia-free. However, as noted by Holland (Holland 1986), the fundamental problem of causal inference is that at most one potential outcome can be realized and thus observed. As a result, to draw inference about the incremental costs of dementia, we must rely on comparisons to dementia-free individuals who may differ along many characteristics. If the distribution of characteristics among individuals with dementia (i.e., cases) is concordant with the distribution of those who are dementia-free (i.e., controls), it is intuitive to believe that differences in their expenditures reveal the incremental cost of dementia. Thus, we begin our estimation of the incremental direct costs of cognitive impairment by conducting a full match of cases (i.e., individuals with undiagnosed dementia and diagnosed dementia) and controls (i.e., individuals without cognitive impairment). We then use an econometric estimation procedure to obtain the incremental costs of cognitive impairment. This approach was used in a recent study to estimate the incremental cost of dementia among individuals diagnosed with dementia (White et al. 2019). Then, we combine our cost estimates with forecasted population counts from the Dementia Population Cost Model to estimate the future economic burden of dementia.

3.2.5.1 Current Cost Estimates

3.2.5.1.1 Matching with a Variable Number of Controls

We used a full matching procedure to control for confounding by creating matched sets containing one HRS respondent with undiagnosed or diagnosed dementia,

with up to five controls (i.e., HRS respondents without cognitive impairment). Each control could be used as a match for up to one case. All HRS respondents used as controls had available Medicare enrollment information over the study period and were matched on birth year, sex, race/ethnicity, education, and HRS entry year. This procedure was conducted to balance the distribution of covariates between controls and respondents with a given level of cognitive impairment so that we could estimate incremental costs of impairment relative to matched individuals without cognitive impairment. We also required that the controls were enrolled in Medicare FFS for the 12 months before and one month after the diagnosis date of their matched case. Because households may pool health care expenditures, we required that neither the control nor any member of their household have cognitive impairment without dementia (CIND), undiagnosed dementia, or diagnosed dementia prior to or for the 72 months following the diagnosis date of their matched case. Control cases were assigned the diagnosis date of their match to equalize time periods

3.2.5.1.2 Econometric Model

We used a two-part model proposed by Basu and Manning (Basu and Manning 2010) to estimate the incremental public expenditure (i.e., Medicare) and out-of-pocket costs for an individual with undiagnosed or diagnosed dementia on a monthly basis. The Basu and Manning estimator extends a prior survival-adjusted estimator (Lin et al. 1997) to appropriately adjust for censoring (e.g., mortality). In the context of the current study, it allows us to account for differential survival between individuals with and without cognitive impairment. In addition, this approach incorporates several well-documented aspects of the cost-accumulation process, such as the increase in costs just prior to death.

Further, this approach decomposes the total marginal cost of cognitive impairment into an intensity of treatment and a survival component. The intensity component, which reflects the marginal cost of cognitive impairment among survivors, corresponds to variation in the accumulation of costs over time among patients with a given level of cognitive impairment compared to those who do not have cognitive impairment. The survival component represents the incremental cost of cognitive impairment associated with cognitive impairment-specific mortality. Individuals with cognitive impairment face increased mortality risk; thus, understanding this survival effect is critical. Further, as life expectancy continues to rise, quantifying the survival effect is increasingly important to understand how costs could vary in response to changes in cognitive impairment-specific mortality. Together, the intensity and survival components shape the total expenditure attributable to cognitive impairment by severity.

Following Basu and Manning (Basu and Manning 2010), we used a three-step algorithm to estimate the mean cost of undiagnosed and diagnosed dementia each as the aggregate of three conditional cost-estimators. Costs were estimated over standardized intervals which we defined on a monthly basis. Monthly costs were scaled up to annual costs and estimated separately for each of the three impairment categories as well as for the first five years since time of diagnosis.

Step 1. Estimate costs in the time intervals prior to death. The first cost-estimator was estimated in two parts. Part 1 only considered time intervals that were observed in their entirety (i.e., in which there was no censoring). That is, intervals in which the individual remained alive; if an individual died on the 15th of a given month, for example, this

interval would not be included in this step. Standard errors were clustered at the individual-level.

$$E(Y_i) = P(Y_i = 1)$$

$$\log \frac{E(Y_i)}{1-E(Y_i)} = \beta_0 + \beta_1 DDem_i + \beta_2 X_i,$$

where Y_i is a binary indicator for whether the costs are greater than zero for individual i, $DDem_i$ is an indicator for whether individual i had a clinical diagnosis of dementia, and X_i is a vector of covariates for subject i. This process was also repeated for undiagnosed dementia by replacing the indicator for diagnosed dementia with an indicator for undiagnosed dementia.

Part 2 of the model only considered intervals that were observed but, unlike Part 1, was restricted to time intervals in which costs were greater than zero. Standard errors were clustered at the individual-level.

$$(E(Y_i|X))^{0.95} = \beta_0 + \beta_1 DDem_i + \beta_2 X_i + \varepsilon_i,$$

where Y_i represents the total costs in the time interval for individual i, $DDem_i$ is an indicator for whether individual i had a clinical diagnosis of dementia, X_i is a vector of covariates for subject i, and ε_i is the error term. Similarly, this process was repeated for

undiagnosed dementia by replacing the indicator for diagnosed dementia with an indicator for undiagnosed dementia.

Step 2. Estimate costs in the intervals of death. We restricted the person-period file to only include intervals during which an individual died. We estimated a model with robust standard errors of the form:

$$(E(Y_i|X))^{0.50} = \beta_0 + \beta_1 DDem_i + \beta_2 X_i + \varepsilon_i,$$

where Y_i corresponds to the total costs in the time interval for individual i, $DDem_i$ is an indicator for whether individual i had a clinical diagnosis of dementia, X_i is a vector of covariates for subject i, and ε_i is the error term. This process was repeated for undiagnosed dementia by replacing the indicator for diagnosed dementia with an indicator for undiagnosed dementia.

Step 3. Estimate the survival function. We used an accelerated failure time model based on the lognormal distribution for time to estimate each individual's survival function while accounting for censoring. This model took the form:

$$\log(T_i) = \beta_0 + \beta_1 DDem + \beta_2 X_i + \varepsilon_i, \varepsilon_i \sim iid\ \mathrm{N}(0,1),$$

where T_i corresponds to the survival time for individual i, $DDem_i$ is an indicator for whether individual i had a clinical diagnosis of dementia, X_i is a vector of covariates for

subject i, and ε_i is the error term. This process was repeated for undiagnosed dementia by replacing the indicator for diagnosed dementia with an indicator for undiagnosed dementia.

In accordance with Basu and Manning, the estimated cost for any individual at time interval j can be specified as

$$\hat{\mu}_j(X) = \hat{S}_j(X) \times \left[\hat{h}_j(X) \times \hat{\mu}_{1j}(X) + \left(1 - \hat{h}_j(X)\right) \times \hat{\mu}_{2j}(X)\right], X = \{D, W\}\ and$$

D is an indicator for impairment and W represents a vector of covariates.

In this equation, $\hat{S}_j(X)$ represents the survival function, $\hat{h}_j(X)$ represents the hazard of death during interval j conditional on an individual's survival until that interval, $\hat{\mu}_{1j}$ represents the estimated cost for an individual if that individual dies in the interval j and $\hat{\mu}_{2j}$ represents what would have been the cost for that individual had they not died in interval j. By partitioning the costs into $\hat{\mu}_{1j}$ and $\hat{\mu}_{2j}$, this equation allows for variation in the accumulation of costs during the interval in which an individual dies compared to those in which they survive.

The marginal effect of dementia on overall cost can then be partitioned into the survival (first term) and intensity (second term) components.

$$\frac{\Delta\mu}{\Delta t} =$$

$$\sum_{j=1}^{K}\left[\left\{\frac{\Delta S_j}{\Delta t}\left[\left(h_j \times \mu_{1j}\right) + \left(\left(1 - h_j\right) \times \mu_{2j}\right)\right] + S_j \times \left[\frac{\Delta h_j}{\Delta t} \times \left(\mu_{1j} - \mu_{2j}\right)\right]\right\} + \left\{S_j \times \left[h_j \times \frac{\Delta\mu_{1j}}{\Delta t} + (1 - h_j) \times \frac{\Delta\mu_{2j}}{\Delta t}\right]\right\}\right]$$

The discrete time modeling approach allows us to specify the intervals of time which we wish to estimate; in this case, months which are aggregated to years. Deltas, denoted by Δ, indicate we are estimating the marginal effect over each time interval. Marginal effects themselves are estimated using the method of recycled predictions (Basu and Rathouz 2005). This method predicts the counterfactual cost for each individual by switching the indicator for a given level of cognitive impairment on and off. Thus, for a given respondent-month, we are able to obtain the actual predicted costs given their level of cognitive impairment and the costs we would have observed if they did not have any cognitive impairment. The difference between the actual and counterfactual estimates provides the marginal effect of exposure (i.e., severity of cognitive impairment) on the survival function, denoted by S_j, and the hazard of death, denoted by h_j.

All models were adjusted for age, sex, race/ethnicity, educational attainment, marital status, quartile of total expenditure for the 12 months prior to dementia diagnosis, and comorbid conditions which included anemia, arthritis, atrial fibrillation, chronic kidney disease, chronic obstructive pulmonary disease, depression, diabetes, heart failure, hypertension, ischemic heart disease, and stroke. To allow for non-linear relationships

between time and costs as well as flexibility with respect to the expenditures by year since and year of diagnosis, models were also adjusted for time since dementia onset (in months) and its interaction with dementia status, indicators for years from diagnosis, and interactions between time from diagnosis and the yearly indicator variables. Standard errors were obtained via bootstrapping with 1,000 iterations.

3.2.5.2 Future Annual Cost Estimates

The number of adults with undiagnosed and dementia disaggregated by age, sex, race/ethnicity, and educational attainment for the years 2000 to 2050 was obtained from the Dementia Population Cost Model presented in Chapter 2. The total annual costs over this period were estimated by multiplying the subgroup- and time-specific costs per individual with the number of individuals in each of the three impairment categories (i.e., undiagnosed dementia, diagnosed dementia). Incremental costs by demographic characteristics and medical comorbidities were held constant; these estimates were adjusted for inflation to 2010 US dollars using the medical care price index (CPI) in the US to account for increases in cost over time.

3.3 Results

3.3.1 Average Cost Estimation per Individual

The period-specific incremental Medicare expenditures and out-of-pocket costs for adults with undiagnosed and diagnosed dementia are shown in Table 3.1. All costs are shown in 2010 dollars. Over a five-year period from dementia onset, the costs to Medicare of undiagnosed and diagnosed dementia relative to persons without cognitive impairment were $1,529 (95%CI: -$3,118, $5,536) and $24,026 (95%CI: $19,865), $28487), respectively. The average out-of-pocket estimates for this same time period

were $6,571 (95%CI: -$38, $14,331) and $28,921 (95%CI: $23,122, $34,551) for undiagnosed and diagnosed dementia.

The HRS sample was sufficiently large to examine costs separately for men and women (Table 3.2). In general, the total incremental costs to Medicare of undiagnosed dementia (men: -$1,564 [95%CI: -$8,999, $6,582]; women: -$2,640 [95%CI: -$13,426, $9,689]) as well as the out-of-pocket costs for undiagnosed (men: $4,955 [95%CI: $1,963, $8,356]; women: $6,459 [95%CI: $2,158, $10,844]) and diagnosed dementia (men: $28,697 [95%CI: $22,445, $36,054]; women: $28,645 [95%CI: $22,599, $36,294]) were comparable as indicated by their overlapping confidence intervals. More substantial differences were observed in the incremental costs to Medicare between men and women for diagnosed dementia men: $24,284 [95%CI: $19,481, $29,190]; women: $1,133 [95%CI: -$2,484, $4,987].

3.3.2 Total Annual Costs

Total current and forecasted incremental costs of undiagnosed and diagnosed dementia are shown in Table 3.3. These estimates were obtained by combining the total incremental costs with forecasted population counts by severity of impairment from the Dementia Population Cost Model in Chapter 2. Thus, the total cost estimate depends on both the cost by severity of impairment and the absolute number of adults with a given level of impairment. The point estimate corresponds to the average cost per person—disaggregated by age, sex, race, educational attainment, and time since diagnosis—combined with the mean forecasted population count for the corresponding subgroup per year from 2000 to 2050. The confidence estimates were obtained by applying the 95% confidence intervals for the subgroup-specific average cost per person to the subgroup-

specific mean forecasted population count. All costs are shown in 2010 dollars to account for inflation. The combined total incremental cost of undiagnosed and diagnosed dementia in 2010 was $145.12 billion (95%CI: $122.15, $176.72). Costs are expected to more than double by 2050 to $378.98 billion (95%CI: $331.03, $446.35) in response to demographic and epidemiological changes in the population. The estimated cost of dementia to Medicare in 2050 is $154.6 billion (95%CI: $131.67, $172.36), an approximate three-fold increase compared to the cost in 2010 of $50.06 billion (95%CI: $40.54, $60.25). Similarly, over this 40-year period, out-of-pocket expenses are expected to rise from $95.06 billion (95%CI: $78.65, $113.45) in 2010 to $224.38 billion (95%CI: $191.59, $252.08) in 2050. These trajectories are illustrated for the period 2000 to 2050 in Figures 3.1 which shows the combined costs for undiagnosed and diagnosed dementia overall.

3.4 Discussion

We report contemporary estimates for the economic burden of dementia in the US using forecasts of cognitive impairment and dementia that account for demographic and epidemiological forces in the population and cost estimates that allow us to evaluate health care costs in the context of differential mortality and intensity of health service utilization at the end of life. Results obtained from our analysis suggest that the direct, incremental out-of-pocket expenditures and costs to Medicare may rise from $145.12 billion (95%CI: $122.15, $176.72) in 2010 to $378.98 billion (95%CI: $331.03, $446.35) by 2050, of which $156.6 (95%CI: 131.57, 172.36) is expected to be covered by Medicare FFS. Importantly, these costs correspond to the incremental or marginal costs associated with dementia rather than the costs of dementia care per se.

Our forecasted, incremental direct costs of dementia are notably higher than recent estimates (Hurd et al. 2013; Hurd et al. 2015; Zissimopoulos et al. 2014). This is partly attributable to differences in the methods used to forecast the number of adults with dementia and partly attributable to differences in the methods used to calculated incremental costs. One recent study (Zissimopoulos et al. 2014) combined total medical cost estimates—for Alzheimer's disease—calculated using the 2002-2004 Medical Expenditure Panel Survey and Medicare Current Beneficiary Survey with the projected number of adults aged 70 years or older with dementia assessed via survey. The authors estimated annual costs per person with Alzheimer's disease of $15,408 in 2010, and forecasted costs which were then multiplied by the size of the population with dementia. Despite the utility of estimates derived from this work, as suggested by the current study and prior research (White et al. 2019), the costs of dementia are known to vary by time since diagnosis and it is possible that survey based measures undercount the true number of dementia cases in the population. In addition, the authors used cost estimated for AD-dementia which represents a majority of dementia cases but may not capture variation across the different types of dementia. Thus, the current approach that estimated costs derived from the same population used to forecast the future number of adults with undiagnosed and diagnosed dementia may provide more accurate estimates of the current and future direct economic burden of dementia.

The extant literature on the cross-sectional and projected costs of dementia illustrates the challenges that arise when drawing comparisons across studies as highlighted in two recent reviews (Fishman et al. 2019; Schaller et al. 2015). Prior work using either administrative (e.g., medical claims) or survey data is limited in that medical insurance

claims may not fully capture the direct costs of dementia including, for example, costs borne by the patient or other insurers. Self-reported measures of expenditure obtained through survey data are also limited as they can be subjective and challenging to standardize across individuals. In addition, prior work has varied with respect to its focus on either social costs (e.g., informal care), economic costs, or a mix of the two. Further, it is likely that the direct costs of dementia accrue prior to the onset which may not be captured in existing cost estimates. Aside from the estimation of costs, prior work has also varied widely in the time period and population studied, the way in which persons with dementia were classified including severity, and the way in which dementia onset was defined. Together with the different ways in which studies have estimated the current or projected size of the population with dementia, direct comparisons of costs are not easily facilitated.

We note several potential weaknesses of our study. First, our study was limited to direct, incremental costs of cognitive impairment rather than comprehensive care costs. The provision of informal care accounts for, in some cases, nearly half or a majority of the total costs of dementia (Hurd et al. 2013). In addition, our estimation of public expenditures was restricted to Medicare FFS (Parts A and B) insurance coverage which reflects a fraction of total public costs for cognitive impairment. Complete claims data for enrollees in Medicare Part C (i.e., Medicare Advantage) are not available and thus could not be included. Data on current Medicaid enrollment was available. However, reliably forecasting future costs borne by Medicaid would require forecasting income distribution, savings, and spending behavior among an aging population to predict who would be eligible for Medicaid in the future. This was not a part of the Dementia Population Cost

Model and thus was excluded from the cost estimates. In addition, although we account for enrollment in private insurance and the corresponding premiums in the out-of-pocket cost estimates, we were not able to obtain reimbursements. Due to these exclusions, our costs may serve as a lower bound for the absolute public expenditure for the incremental direct cost of medical care (excluding long-term care services). Second, because long-term care services are not covered by Medicare beyond the first 100 days, we were not able to count for differential use of these services between cases and controls. Third, the accuracy of self-reported out-of-pocket expenditures was not assessed, and these estimates could be imprecise. Fourth, due to data constraints, we were not able to estimate annual costs beyond the fifth year of cognitive impairment for any level of severity; thus, we assumed constancy of costs after the fifth year. Finally, the forecasted population counts by severity of cognitive impairment were obtained from the Dementia Population Cost Model which is constrained to its own set of limitations. Despite these potential weaknesses, our study provides important insight regarding the costs of undiagnosed to diagnosed dementia from incident impairment to death.

Our study also had several notable strengths. We used a large, nationally representative and longitudinal sample of older adults with linked Medicare claims data. This allowed us to combine survey-based measures of out-of-pocket expenses with validated information on health service costs and utilization from a period prior to onset and until death. Whereas prior studies focused primarily on either survey- or claims-based measures of dementia, we estimated costs separately for adults with undiagnosed dementia and diagnosed. Our inclusion of both undiagnosed and diagnosed dementia cases overcomes limitations in prior work that either rely on survey-based measures or

administrative data which may underestimate the economic burden of dementia among adults whose cognitive impairment is not yet recognized by their healthcare system which itself may be a reflection of their access to healthcare or how they utilize services. In addition, because of the wide range of characteristics assessed in the HRS, we were able to adjusted for important sociodemographic and health-related characteristics important to the estimation of costs which are typically unavailable in claims databases. For our estimation of costs, we used a full matching procedure to balance covariates between individuals with and without cognitive impairment. In addition, we used an innovative econometric approach that allowed us to account for differential mortality by severity of cognitive impairment as well as variation in health service utilization which both contribute to the estimation of health care costs. Finally, our forecasted cost estimates were based on the Dementia Population Cost Model which, despite its limitations, accounts for demographic and epidemiological changes in the population.

Dementia is the most expensive health condition in the United States. In addition to increased health care utilization resulting in higher out-of-pocket and public expenditures, individuals with dementia face increased direct costs associated with formal home care as well as long-term care services. Despite this study's focus on direct, incremental costs, the economic and social burden of dementia extend far beyond and weigh heavily on other economic and social care systems, underscoring the importance of addressing the burden of dementia from economic, health systems, and social care financing perspectives. The current study illustrates how the direct costs of dementia vary by whether or not an individual has a diagnostic claim, how these costs are expected to change over time, and how this burden is distributed among different payers. We found

substantial differences in direct costs among adults with undiagnosed dementia and diagnosed dementia. Whether this is due to differences in access to or utilization of health services is an important line of inquiry that may help shape future policies that address the economic burden of dementia among population subgroups.

The social and economic burden of dementia will change in the coming decades as we achieve gains in life expectancy. Improved survival with dementia could increase overall costs on the basis of higher duration lived with the condition. A delay in the onset of dementia would be beneficial in reducing the indirect costs of care, but would not eradicate the direct costs in the first year after diagnosis which exceeds direct costs in all successive years. Thus, in the absence of a cure, improvements in population health and medical advances may not necessarily yield reductions in health care expenditures.

Federal investment in Alzheimer’s disease research has increased since the passage of the National Alzheimer's Project Act (NAPA) and the Alzheimer's Accountability Act in 2011 (Health and Services 2012). Despite historic investments in AD research funding since that time, including a record $2.8 billion in 2020 (Alzheimer's Association 2020), these funds still represent a fraction of the direct economic burden. The substantial growth of Medicare costs in the coming decades merit a comprehensive review to improve care and cost-effectiveness of care.

3.5 Tables

Table 3.1 Period-specific incremental Medicare expenditures and out-of-pocket costs for adults with undiagnosed and dementia, HRS 1992-2012

	Medicare		Out-of-pocket	
Period	Undiagnosed dementia	Diagnosed dementia	Undiagnosed dementia	Diagnosed dementia
Year 1	618 (-837, 2392)	7524 (6613,8636)	475 (-540, 1413)	3003 (2255, 3921)
Year 2	178 (-970, 1784)	5976 (5273,6818)	921 (-49, 1916)	4586 (3833, 5299)
Year 3	-195 (-1406, 1192)	4738 (4030,5447)	1362 (210, 2673)	5944 (5009, 6951)
Year 4	-573 (-2090, 1181)	3502 (2695,4310)	1771 (256, 3718)	7152 (5720, 8513)
Year 5	-1021 (-3307, 1202)	2285 (1254,3276)	2042 (85, 4611)	8237 (6304, 9868)
Total	1529 (-3118, 5536)	24026 (19865, 28487)	6571 (-38, 14331)	28921 (23122, 34551)

Notes. Costs are shown in 2010 dollars. Data in parentheses are 95% confidence intervals.

Table 3.2 Period-specific incremental Medicare expenditures and out-of-pocket costs for adults with undiagnosed and dementia for men and women, HRS 1992-2012

	Medicare		Out-of-pocket	
Men	Undiagnosed dementia	Diagnosed dementia	Undiagnosed dementia	Diagnosed dementia
Year 1	585 (-817, 1959)	7703 (6572, 8923)	568 (83, 1061)	2887 (2110, 4070)
Year 2	101 (-1170, 1377)	6088 (5288, 6958)	821 (360, 1307)	4530 (3843, 5315)
Year 3	-313 (-1673, 1070)	4788 (4037, 5535)	1022 (505, 1662)	5921 (4804, 7284)
Year 4	-722 (-2311, 1100)	3492 (2571, 4407)	1233 (549, 2101)	7138 (5560, 8842)
Year 5	-1215 (-3028, 1077)	2212 (1013, 3367)	1310 (466, 2226)	8221 (6128, 10543)
Total	-1564 (-8999, 6582)	24284 (19481, 29190)	4955 (1963, 8356)	28697 (22445, 36054)
Women				
Year 1	1130 (-829, 3339)	-590 (-1232, 20)	798 (181, 1461)	2964 (2263, 3808)
Year 2	231 (-1307, 2207)	-214 (-800, 365)	1084 (589, 1739)	4528 (3716, 5176)
Year 3	-527 (-2153, 1777)	197 (-394, 840)	1326 (602, 1950)	5874 (4754, 7328)
Year 4	-1287 (-3701, 1392)	644 (-125, 1490)	1594 (488, 2698)	7088 (5587, 9177)
Year 5	-2187 (-5435, 974)	1096 (67, 2272)	1658 (298, 2995)	8191 (6280, 10806)
Total	-2640 (-13426, 9689)	1133 (-2484, 4987)	6459 (2158, 10844)	28645 (22599, 36294)

Notes. Costs are shown in 2010 dollars. Data in parentheses are 95% confidence intervals.

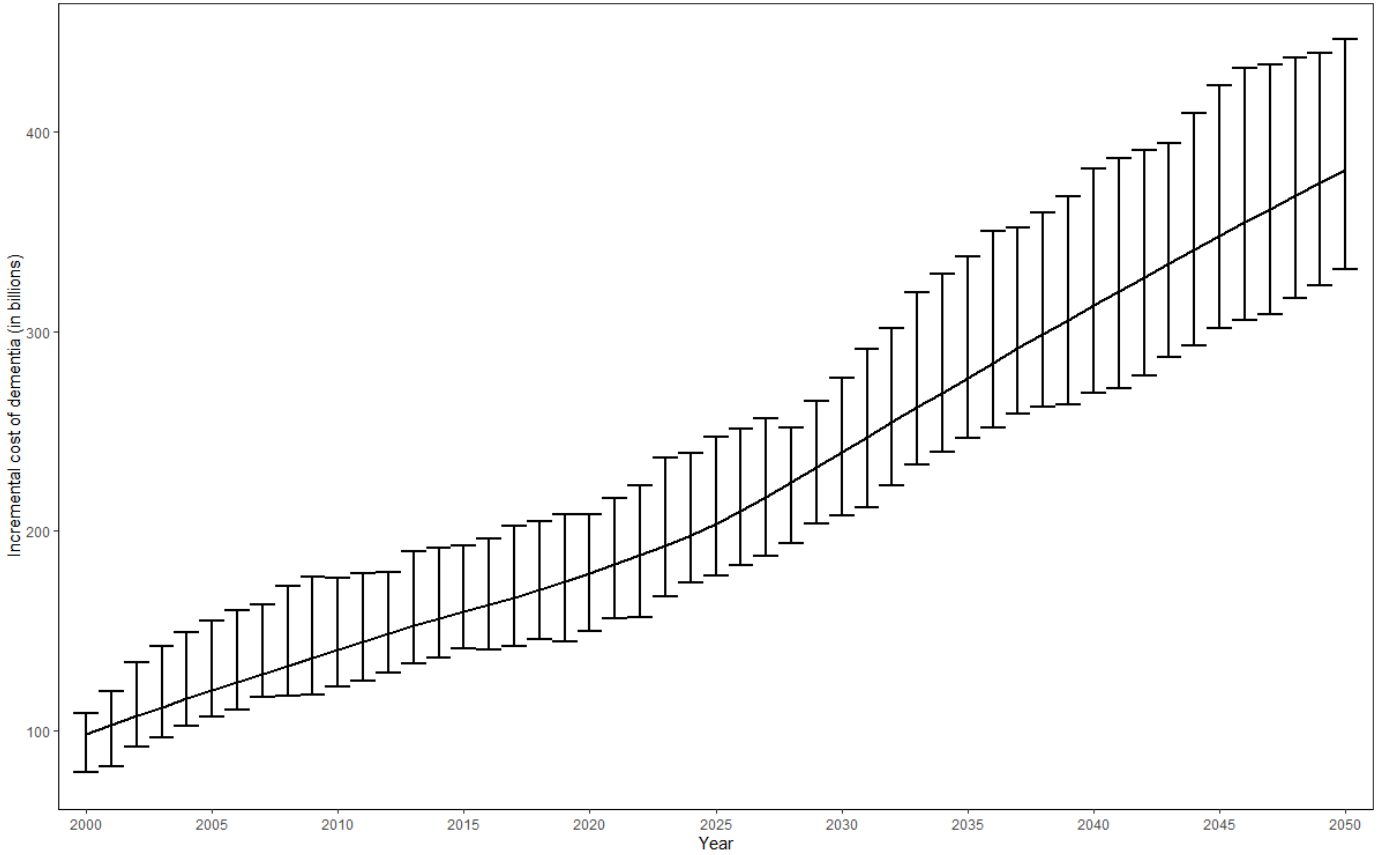
Table 3.3 Forecasted total incremental costs of undiagnosed dementia and diagnosed dementia overall and by out-of-pocket costs and Medicare expenditures

	Cost (in billions), Overall		
Year	Total	Medicare	Out-of-pocket
2010	145.12 (122.15, 176.72)	50.06 (40.54, 60.35)	95.06 (78.56, 113.45)
2020	177.25 (149.62, 208.27)	66.64 (54.76, 76.35)	110.61 (90.42, 126.73)
2030	236.74 (208.02, 276.75)	92.35 (81.01, 103.87)	144.39 (123.08, 165.68)
2040	312.29 (269.5, 381.9)	135.23 (115.54, 157.75)	177.06 (150.21, 209.38)
2050	378.98 (331.03, 446.35)	154.6 (131.67, 172.36)	224.38 (191.59, 252.08)

Notes. Costs are shown in 2010 dollars. Data in parentheses are 95% confidence intervals for cost.

3.6 Figures

Figure 3.1 Forecasted total incremental out-of-pocket and Medicare costs of undiagnosed dementia and diagnosed dementia, 2000-2050



Notes. Vertical bars correspond to 95% confidence intervals for cost.

REFERENCES

Alley, D., K. Suthers, and E. Crimmins. 2007. "Education and Cognitive Decline in Older Americans: Results From the AHEAD Sample." *Res Aging* 29(1):73-94.

Altmann, A., L. Tian, V.W. Henderson, M.D. Greicius, and I. Alzheimer's Disease Neuroimaging Initiative. 2014. "Sex modifies the APOE-related risk of developing Alzheimer disease." *Ann Neurol* 75(4):563-573.

Alzheimer's Association. 2019. "2019 Alzheimer's disease facts and figures." *Alzheimers & Dementia* 15(3):321-387.

Alzheimer's Association. 2019. "Costs of Alzheimer's to Medicare and Medicaid."

—. 2020. "Fiscal Year 2021 Alzheimer's Research Funding."

American Psychiatric Association. 1980. "Diagnostic and statistical manual of mental disorders; DSM-III." Washington, DC.

—. 1994. "Diagnostic and statistical manual of mental disorders; DSM-IV." Washington, DC.

Anjum, I., M. Fayyaz, A. Wajid, W. Sohail, and A. Ali. 2018. "Does Obesity Increase the Risk of Dementia: A Literature Review." *Cureus* 10(5):e2660.

Anstey, K.J., C. von Sanden, A. Salim, and R. O'Kearney. 2007. "Smoking as a risk factor for dementia and cognitive decline: a meta-analysis of prospective studies." *American journal of epidemiology* 166(4):367-378.

Arroyo-Johnson, C.and K.D. Mincey. 2016. "Obesity Epidemiology Worldwide." *Gastroenterology clinics of North America* 45(4):571-579.

Babulal, G.M., Y.T. Quiroz, B.C. Albensi, E. Arenaza-Urquijo, A.J. Astell, C. Babiloni, A. Bahar-Fuchs, J. Bell, G.L. Bowman, A.M. Brickman, G. Chetelat, C. Ciro, A.D. Cohen, P. Dilworth-Anderson, H.H. Dodge, S. Dreux, S. Edland, A. Esbensen, L. Evered, M. Ewers, K.N. Fargo, J. Fortea, H. Gonzalez, D.R. Gustafson, E. Head, J.A. Hendrix, S.M. Hofer, L.A. Johnson, R. Jutten, K. Kilborn, K.L. Lanctot, J.J. Manly, R.N. Martins, M.M. Mielke, M.C. Morris, M.E. Murray, E.S. Oh, M.A. Parra, R.A. Rissman, C.M. Roe, O.A. Santos, N. Scarmeas, L.S. Schneider, N. Schupf, S. Sikkes, H.M. Snyder, H.R. Sohrabi, Y. Stern, A. Strydom, Y. Tang, G.M. Terrera, C. Teunissen, D. Melo van Lent, M. Weinborn, L. Wesselman, D.M. Wilcock, H. Zetterberg, S.E. O'Bryant, R. International Society to Advance Alzheimer's, and A.s.A. Treatment. 2019. "Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: Update and areas of immediate need." *Alzheimers Dement* 15(2):292-312.

—. 2019. "Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: Update and areas of immediate need." *Alzheimers Dement* 15(2):292-312.

Bäckman, L., B.J. Small, and A. Wahlin. 2001. "Aging and memory." *Handbook of the psychology of aging*:349-377.

Barbieri, M., J.R. Wilmoth, V.M. Shkolnikov, D. Glei, D. Jasilionis, D. Jdanov, C. Boe, T. Riffe, P. Grigoriev, and C. Winant. 2015. "Data Resource Profile: The Human Mortality Database (HMD)." *International journal of epidemiology* 44(5):1549-1556.

Barnes, D.and K. Yaffe. 2011. "Lancet Neurol." *The projected effect of risk factor reduction on Alzheimer's disease prevalence* 10(9):819-828.

Barnes, D.E.and K. Yaffe. 2011. "The projected effect of risk factor reduction on Alzheimer's disease prevalence." *Lancet Neurol* 10(9):819-828.

Barnes, L.L.and D.A. Bennett. 2014. "Alzheimer's disease in African Americans: risk factors and challenges for the future." *Health affairs (Project Hope)* 33(4):580-586.

Basu, A.and W.G. Manning. 2010. "Estimating lifetime or episode-of-illness costs under censoring." *Health Econ* 19(9):1010-1028.

Basu, A.and P.J. Rathouz. 2005. "Estimating marginal and incremental effects on health outcomes using flexible link and variance function models." *Biostatistics* 6(1):93-109.

Benjamin, E.J., M.J. Blaha, S.E. Chiuve, M. Cushman, S.R. Das, R. Deo, S.D. de Ferranti, J. Floyd, M. Fornage, C. Gillespie, C.R. Isasi, M.C. Jimenez, L.C. Jordan, S.E. Judd, D. Lackland, J.H. Lichtman, L. Lisabeth, S. Liu, C.T. Longenecker, R.H. Mackey, K. Matsushita, D. Mozaffarian, M.E. Mussolino, K. Nasir, R.W. Neumar, L. Palaniappan, D.K. Pandey, R.R. Thiagarajan, M.J. Reeves, M. Ritchey, C.J. Rodriguez, G.A. Roth, W.D. Rosamond, C. Sasson, A. Towfighi, C.W. Tsao, M.B. Turner, S.S. Virani, J.H. Voeks, J.Z. Willey, J.T. Wilkins, J.H. Wu, H.M. Alger, S.S. Wong, P. Muntner, C. American Heart Association Statistics, and S. Stroke Statistics. 2017. "Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association." *circulation* 135(10):e146-e603.

Berry, W.D. 1993. *Understanding regression assumptions*: Sage Publications.

Betancourt, J.R., A.R. Green, and J.E. Carrillo. 2002. *Cultural competence in health care: Emerging frameworks and practical approaches*: Commonwealth Fund, Quality of Care for Underserved Populations New York, NY.

Bharmal, M.F., S. Dedhiya, B.A. Craig, M. Weiner, M. Rosenman, L.P. Sands, A. Modi, C. Doebbeling, and J. Thomas. 2012. "Incremental dementia-related expenditures in a medicaid population." *Am J Geriatr Psychiatry* 20(1):73-83.

Brandt, J., M. Spencer, and M. Folstein. 1988. "The Telephone Interview for Cognitive Status." *Cognitive and Behavioral Neurology* 1(2):111-118.

Brookmeyer, R., S. Gray, and C. Kawas. 1998. "Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset." *American journal of public health* 88(9):1337-1342.

Buuren, S.v.and K. Groothuis-Oudshoorn. 2010. "mice: Multivariate imputation by chained equations in R." *Journal of statistical software*:1-68.

Byford, S., D.J. Torgerson, and J. Raftery. 2000. "Economic note: cost of illness studies." *Bmj* 320(7245):1335.

Bynum, J.P., P.V. Rabins, W. Weller, M. Niefeld, G.F. Anderson, and A.W. Wu. 2004. "The relationship between a dementia diagnosis, chronic illness, medicare expenditures, and hospital use." *Journal of the American Geriatrics Society* 52(2):187-194.

Centers for Disease Control and Prevention. " Behavioral Risk Factor Surveillance System

Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2008-2018.".

Centers for Medicare & Medicaid Services. 2018. "Accessing CMS Data."

Centers for Medicare Medicaid Services. 2018. *Medicare and You 2018*: US Department of Health and Human Services, Centers for Medicare and Medicaid Services.

Chang, C.-C.H., Y. Zhao, C.-W. Lee, and M. Ganguli. 2012. "Smoking, death, and Alzheimer's disease: a case of competing risks." *Alzheimer disease and associated disorders* 26(4):300.

Chen, Y., B. Tysinger, E. Crimmins, and J.M. Zissimopoulos. 2019. "Analysis of dementia in the US population using Medicare claims: Insights from linked survey and administrative claims data." *Alzheimer's & dementia (New York, N. Y.)* 5:197-207.

Chêne, G., A. Beiser, R. Au, S.R. Preis, P.A. Wolf, C. Dufouil, and S. Seshadri. 2015. "Gender and incidence of dementia in the Framingham Heart Study from mid-adult life." *Alzheimer's & Dementia* 11(3):310-320.

Chhatwal, J., S. Jayasuriya, and E.H. Elbasha. 2016. "Changing Cycle Lengths in State-Transition Models: Challenges and Solutions." *Med Decis Making* 36(8):952-964.

Chin, A.L., S. Negash, and R. Hamilton. 2011. "Diversity and disparity in dementia: the impact of ethnoracial differences in Alzheimer disease." *Alzheimer Dis Assoc Disord* 25(3):187-195.

Coe, N.B., M.M. Skira, and E.B. Larson. 2018. "A Comprehensive Measure of the Costs of Caring for a Parent: Differences According to Functional Status." *J Am Geriatr Soc* 66(10):2003-

2008.

Colby, S.L.and J.M. Ortman. 2017. "Projections of the size and composition of the US population: 2014 to 2060: Population estimates and projections."

Cole, S.R.and C.E. Frangakis. 2009. "The consistency statement in causal inference: a definition or an assumption?" *Epidemiology* 20(1):3-5.

Cole, S.R., M.G. Hudgens, M.A. Brookhart, and D. Westreich. 2015. "Risk." *Am J Epidemiol* 181(4):246-250.

Cole, S.R., B. Lau, J.J. Eron, M.A. Brookhart, M.M. Kitahata, J.N. Martin, W.C. Mathews, and M.J. Mugavero. 2014. "Estimation of the standardized risk difference and ratio in a competing risks framework: application to injection drug use and progression to AIDS after initiation of antiretroviral therapy." *American journal of epidemiology* 181(4):238-245.

Crimmins, E.M., J.K. Kim, K.M. Langa, and D.R. Weir. 2011. "Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study." *J Gerontol B Psychol Sci Soc Sci* 66 Suppl 1:i162-171.

Crimmins, E.M., Y. Saito, J.K. Kim, Y.S. Zhang, I. Sasson, and M.D. Hayward. 2018. "Educational Differences in the Prevalence of Dementia and Life Expectancy with Dementia: Changes from 2000 to 2010." *J Gerontol B Psychol Sci Soc Sci* 73(suppl_1):S20-S28.

Danaei, G., A. Pan, F.B. Hu, and M.A. Hernán. 2013. "Hypothetical mid-life interventions in women and risk of type 2 diabetes." *Epidemiology (Cambridge, Mass.)* 24(1):122.

Day, J.C.and K. Bauman. 2000. *Have We Reached the Top?: Educational Attainment Projections of the US Population*: Population Division, US Census Bureau.

Deb, A., J.D. Thornton, U. Sambamoorthi, and K. Innes. 2017. "Direct and indirect cost of managing Alzheimer's disease and related dementias in the United States." *Expert review of pharmacoeconomics & outcomes research* 17(2):189-202.

Denkinger, M.D., T. Nikolaus, C. Denkinger, and A. Lukas. 2012. "Physical activity for the prevention of cognitive decline: current evidence from observational and controlled studies." *Z Gerontol Geriatr* 45(1):11-16.

Derby, C.A., M.J. Katz, R.B. Lipton, and C.B. Hall. 2017. "Trends in Dementia Incidence in a Birth Cohort Analysis of the Einstein Aging Study." *JAMA Neurol* 74(11):1345-1351.

Doblhammer, G., A. Fink, S. Zylla, T. Fritze, and F. Willekens. 2014. "Short-term trends in German dementia prevalence, incidence, and mortality." *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 10(4):P279.

Durazzo, T.C., N. Mattsson, M.W. Weiner, and I. Alzheimer's Disease Neuroimaging. 2014. "Smoking and increased Alzheimer's disease risk: a review of potential mechanisms." *Alzheimers Dement* 10(3 Suppl):S122-145.

Efron, B.and R. Tibshirani. 1986. "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy." *Statistical science*:54-75.

Eskelinen, M.H., T. Ngandu, J. Tuomilehto, H. Soininen, and M. Kivipelto. 2009. "Midlife

coffee and tea drinking and the risk of late-life dementia: a population-based CAIDE study." *J Alzheimers Dis* 16(1):85-91.

Exalto, L.G., C.P. Quesenberry, D. Barnes, M. Kivipelto, G.J. Biessels, and R.A. Whitmer. 2014. "Midlife risk score for the prediction of dementia four decades later." *Alzheimers Dement* 10(5):562-570.

Farina, N., J. Rusted, and N. Tabet. 2014. "The effect of exercise interventions on cognitive outcome in Alzheimer's disease: a systematic review." *Int Psychogeriatr* 26(1):9-18.

Feder, J., H.L. Komisar, and M. Niefeld. 2000. "Long-term care in the United States: an overview." *Health affairs (Project Hope)* 19(3):40-56.

Fernando, M.S., P.G. Ince, M.R.C.C. Function, and G. Ageing Neuropathology Study. 2004. "Vascular pathologies and cognition in a population-based cohort of elderly people." *J Neurol Sci* 226(1-2):13-17.

Fine, J.P.and R.J. Gray. 1999. "A proportional hazards model for the subdistribution of a competing risk." *Journal of the American statistical association* 94(446):496-509.

Fisher, G., H. Hassan, W. Rodgers, and D. Weir. 2013. "Health and retirement study imputation of cognitive functioning measures: 1992–2010." *Journal of Occupational Health Psychology* 19:231-242.

Fisher, G.G.and L.H. Ryan. 2018. "Overview of the Health and Retirement Study and Introduction to the Special Issue." *Work Aging Retire* 4(1):1-9.

Fishman, P., N.B. Coe, L. White, P.K. Crane, S. Park, B. Ingraham, and E.B. Larson. 2019. "Currently Reading Cost of Dementia in Medicare Managed Care: A Systematic Literature Review." *The American journal of managed care*.

Fitzpatrick, A.L., L.H. Kuller, O.L. Lopez, P. Diehr, E.S. O'Meara, W.T. Longstreth, Jr., and J.A. Luchsinger. 2009. "Midlife and late-life obesity and the risk of dementia: cardiovascular health study." *Arch Neurol* 66(3):336-342.

Flegal, K.M., M.D. Carroll, B.K. Kit, and C.L. Ogden. 2012. "Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010." *Jama* 307(5):491-497.

Fletcher, B., M. Gulanick, and C. Lamendola. 2002. "Risk factors for type 2 diabetes mellitus." *J Cardiovasc Nurs* 16(2):17-23.

Folstein, M.F., S.E. Folstein, and P.R. McHugh. 1975. "“Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician." *Journal of psychiatric research* 12(3):189-198.

Force, U.P.S.T. 2010. "Screening for obesity in children and adolescents: US Preventive Services Task Force recommendation statement." *Pediatrics* 125(2):361-367.

Freedman, V.A., B.C. Spillman, P.M. Andreski, J.C. Cornman, E.M. Crimmins, E. Kramarow, J. Lubitz, L.G. Martin, S.S. Merkin, R.F. Schoeni, T.E. Seeman, and T.A. Waidmann. 2013. "Trends in late-life activity limitations in the United States: an update from five national surveys." *Demography* 50(2):661-671.

Friedman, E.M., R.A. Shih, K.M. Langa, and M.D. Hurd. 2015. "US Prevalence And Predictors Of Informal Caregiving For Dementia." *Health affairs (Project Hope)* 34(10):1637-1641.

Frytak, J.R., H.J. Henk, Y. Zhao, L. Bowman, J.A. Flynn, and M. Nelson. 2008. "Health service utilization among Alzheimer's disease patients: evidence from managed care." *Alzheimer's & Dementia* 4(5):361-367.

Ganguli, M. 2017. "The times they are a-changin': cohort effects in aging, cognition, and dementia 1." *International psychogeriatrics* 29(3):353-355.

Garcia-Aymerich, J., R. Varraso, G. Danaei, C.A. Camargo, Jr., and M.A. Hernan. 2014. "Incidence of adult-onset asthma after hypothetical interventions on body mass index and physical activity: an application of the parametric g-formula." *Am J Epidemiol* 179(1):20-26.

Geiss, L.S., J. Wang, Y.J. Cheng, T.J. Thompson, L. Barker, Y. Li, A.L. Albright, and E.W. Gregg. 2014. "Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012." *Jama* 312(12):1218-1226.

Gilmer, T.P.and P.J. O'Connor. 2010. "The growing importance of diabetes screening." *Diabetes care* 33(7):1695-1697.

Glied, S. 1996. "Estimating the indirect cost of illness: an assessment of the forgone earnings approach." *American journal of public health* 86(12):1723-1728.

Goldfeld, K.S., D.G. Stevenson, M.B. Hamel, and S.L. Mitchell. 2011. "Medicare expenditures among nursing home residents with advanced dementia." *Archives of internal medicine*

171(9):824-830.

Goldman, D.P., P.G. Shekelle, J. Bhattacharya, M. Hurd, and G.F. Joyce. 2004. "Health status and medical treatment of the future elderly." RAND CORP SANTA MONICA CA.

Gooley, T.A., W. Leisenring, J. Crowley, and B.E. Storer. 1999. "Estimation of failure probabilities in the presence of competing risks: new representations of old estimators." *Stat Med* 18(6):695-706.

Gorina, Y.and E.A. Kramarow. 2011. "Identifying chronic conditions in Medicare claims data: evaluating the Chronic Condition Data Warehouse algorithm." *Health services research* 46(5):1610-1627.

Graves, A., C. Van Duijn, V. Chandra, L. Fratiglioni, A. Heyman, A. Jorm, E. Kokmen, K. Kondo, J. Mortimer, and W.A. Rocca. 1991. "Alcohol and tobacco consumption as risk factors for Alzheimer's disease: a collaborative re-analysis of case-control studies." *International journal of epidemiology* 20(Supplement_2):S48-S57.

Greenland, S. 2003. "Quantifying biases in causal models: classical confounding vs collider-stratification bias." *Epidemiology* 14(3):300-306.

Greenland, S., J. Pearl, and J.M. Robins. 1999. "Causal diagrams for epidemiologic research." *Epidemiology* 10(1):37-48.

Groot, C., A.M. Hooghiemstra, P.G. Raijmakers, B.N. van Berckel, P. Scheltens, E.J. Scherder, W.M. van der Flier, and R. Ossenkoppele. 2016. "The effect of physical activity on cognitive

function in patients with dementia: A meta-analysis of randomized control trials." *Ageing Res Rev* 25:13-23.

Gudala, K., D. Bansal, F. Schifano, and A. Bhansali. 2013. "Diabetes mellitus and risk of dementia: A meta-analysis of prospective observational studies." *J Diabetes Investig* 4(6):640-650.

Gurland, B.J., D.E. Wilder, R. Lantigua, Y. Stern, J. Chen, E.H. Killeffer, and R. Mayeux. 1999. "Rates of dementia in three ethnoracial groups." *Int J Geriatr Psychiatry* 14(6):481-493.

Haan, M.N., D.M. Mungas, H.M. Gonzalez, T.A. Ortiz, A. Acharya, and W.J. Jagust. 2003. "Prevalence of dementia in older latinos: the influence of type 2 diabetes mellitus, stroke and genetic factors." *J Am Geriatr Soc* 51(2):169-177.

Hardy, J.A.and G.A. Higgins. 1992. "Alzheimer's disease: the amyloid cascade hypothesis." *Science* 256(5054):184-185.

Health, U.D.o.and H. Services. 2012. "National plan to address Alzheimer’s disease."

Hebert, L. 2000. "Scherr PA, Bienias JL, Bennett DA, Evans DA." *Alzheimer disease in the US population: prevalence estimates using the*:1119-1122.

Hebert, L.E., P.A. Scherr, L.A. Beckett, H.H. Funkenstein, M.S. Albert, M.J. Chown, and D.A. Evans. 1992. "Relation of smoking and alcohol consumption to incident Alzheimer's disease." *Am J Epidemiol* 135(4):347-355.

Hebert, L.E., P.A. Scherr, J.L. Bienias, D.A. Bennett, and D.A. Evans. 2003. "Alzheimer disease

in the US population: prevalence estimates using the 2000 census." *Archives of neurology* 60(8):1119-1122.

Hebert, L.E., J. Weuve, P.A. Scherr, and D.A. Evans. 2013. "Alzheimer disease in the United States (2010–2050) estimated using the 2010 census." *Neurology* 80(19):1778-1783.

Heeringa, S.G.and J.H. Connor. 1995. "Technical description of the Health and Retirement Survey sample design." *Ann Arbor: University of Michigan*.

Heeringa, S.G., G.G. Fisher, M. Hurd, K.M. Langa, M.B. Ofstedal, B.L. Plassman, W.L. Rodgers, and D.R. Weir. 2009. "Aging, Demographics and Memory Study (ADAMS) sample design, weighting and analysis for ADAMS."

Hernán, M.A., S. Hernández-Díaz, and J.M. Robins. 2004. "A structural approach to selection bias." *Epidemiology* 15(5):615-625.

Herzog, A.R.and R.B. Wallace. 1997. "Measures of cognitive functioning in the AHEAD Study." *J Gerontol B Psychol Sci Soc Sci* 52 Spec No(Special_Issue):37-48.

Holland, P.W. 1986. "Statistics and causal inference." *Journal of the American statistical association* 81(396):945-960.

HRS Staff. 2008. "HRS Sample Evolution: 1992-1998."

Hruby, A., J.E. Manson, L. Qi, V.S. Malik, E.B. Rimm, Q. Sun, W.C. Willett, and F.B. Hu. 2016. "Determinants and Consequences of Obesity." *American journal of public health* 106(9):1656-1662.

Human Mortality Database. *URL: http://mortality.org (04 April 2019, date last accessed).*

Hurd, M.D., P. Martorell, A. Delavande, K.J. Mullen, and K.M. Langa. 2013. "Monetary costs of dementia in the United States." *New England Journal of Medicine* 368(14):1326-1334.

Hurd, M.D., P. Martorell, and K. Langa. 2015. "Future Monetary Costs of Dementia in the United States under Alternative Dementia Prevalence Scenarios." *Journal of population ageing* 8(1-2):101-112.

Islam, M.A.and R.I. Chowdhury. 2006. "A higher order Markov model for analyzing covariate dependence." *Applied Mathematical Modelling* 30(6):477-488.

Jefferson, A.L. 2014. "Vascular risk factors and midlife cognition: rethinking the exposure window." *circulation* 129(15):1548-1550.

Jones, D.S.and J.A. Greene. 2016. "Is Dementia in Decline? Historical Trends and Future Trajectories." *N Engl J Med* 374(6):507-509.

Jorm, A.F. 1994. "A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): development and cross-validation." *Psychol Med* 24(1):145-153.

Jorm, A.F.and D. Jolley. 1998. "The incidence of dementia: a meta-analysis." *Neurology* 51(3):728-733.

Joyce, A.T., Y. Zhao, L. Bowman, J.A. Flynn, C.T. Carter, and D.A. Ollendorf. 2007. "Burden of illness among commercially insured patients with Alzheimer’s disease." *Alzheimer's & Dementia* 3(3):204-210.

Jutkowitz, E., R.L. Kane, J.E. Gaugler, R.F. MacLehose, B. Dowd, and K.M. Kuntz. 2017. "Societal and Family Lifetime Cost of Dementia: Implications for Policy." *J Am Geriatr Soc* 65(10):2169-2175.

Kalaria, R.N. 2012. "Cerebrovascular disease and mechanisms of cognitive impairment: evidence from clinicopathological studies in humans." *Stroke* 43(9):2526-2534.

Karlawish, J. 2011. "Addressing the ethical, policy, and social challenges of preclinical Alzheimer disease." *Neurology* 77(15):1487-1493.

Karp, A., I. Kåreholt, C. Qiu, T. Bellander, B. Winblad, and L. Fratiglioni. 2004. "Relation of education and occupation-based socioeconomic status to incident Alzheimer's disease." *American journal of epidemiology* 159(2):175-183.

Kelley, A.S., K. McGarry, R. Gorges, and J.S. Skinner. 2015. "The burden of health care costs for patients with dementia in the last 5 years of life." *Ann Intern Med* 163(10):729-736.

Kernan, W.N., B. Ovbiagele, H.R. Black, D.M. Bravata, M.I. Chimowitz, M.D. Ezekowitz, M.C. Fang, M. Fisher, K.L. Furie, D.V. Heck, S.C. Johnston, S.E. Kasner, S.J. Kittner, P.H. Mitchell, M.W. Rich, D. Richardson, L.H. Schwamm, J.A. Wilson, C.o.C. American Heart Association Stroke Council, C.o.C.C. Stroke Nursing, and D. Council on Peripheral Vascular. 2014. "Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association." *Stroke* 45(7):2160-2236.

Kivimaki, M., R. Luukkonen, G.D. Batty, J.E. Ferrie, J. Pentti, S.T. Nyberg, M.J. Shipley, L.

Alfredsson, E.I. Fransson, M. Goldberg, A. Knutsson, M. Koskenvuo, E. Kuosma, M. Nordin, S.B. Suominen, T. Theorell, E. Vuoksimaa, P. Westerholm, H. Westerlund, M. Zins, M. Kivipelto, J. Vahtera, J. Kaprio, A. Singh-Manoux, and M. Jokela. 2018. "Body mass index and risk of dementia: Analysis of individual-level data from 1.3 million individuals." *Alzheimers Dement* 14(5):601-609.

Kivimaki, M., A. Singh-Manoux, J. Pentti, S. Sabia, S.T. Nyberg, L. Alfredsson, M. Goldberg, A. Knutsson, M. Koskenvuo, A. Koskinen, A. Kouvonen, M. Nordin, T. Oksanen, T. Strandberg, S.B. Suominen, T. Theorell, J. Vahtera, A. Vaananen, M. Virtanen, P. Westerholm, H. Westerlund, M. Zins, S. Seshadri, G.D. Batty, P.N. Sipila, M.J. Shipley, J.V. Lindbohm, J.E. Ferrie, M. Jokela, and I.P.-W. consortium. 2019. "Physical inactivity, cardiometabolic disease, and risk of dementia: an individual-participant meta-analysis." *Bmj* 365:l1495.

Kochanek, K.D., S.L. Murphy, J. Xu, and B. Tejada-Vera. 2016. "Deaths: Final Data for 2014." *Natl Vital Stat Rep* 65(4):1-122.

Kontari, P.and K.J. Smith. 2019. "Risk of dementia associated with cardiometabolic abnormalities and depressive symptoms: a longitudinal cohort study using the English longitudinal study of ageing." *Int J Geriatr Psychiatry* 34(2):289-298.

Kramarow, E.A.and B. Tejada-Vera. 2019. "National Vital Statistics Reports." *National Vital Statistics Reports* 68(2).

Kuo, T.C., Y. Zhao, S. Weir, M.S. Kramer, and A.S. Ash. 2008. "Implications of comorbidity on costs for patients with Alzheimer disease." *Med Care* 46(8):839-846.

Langa, K., M. Kabeto, and D. Weir. 2010. "Report on race and cognitive impairment using HRS in, 2010 Alzheimer's disease facts and figures 2010." *Alzheimer's Disease Facts and Figures*.

Langa, K.M., E.B. Larson, E.M. Crimmins, J.D. Faul, D.A. Levine, M.U. Kabeto, and D.R. Weir. 2017. "A Comparison of the Prevalence of Dementia in the United States in 2000 and 2012." *JAMA Intern Med* 177(1):51-58.

Langa, K.M., E.B. Larson, J.H. Karlawish, D.M. Cutler, M.U. Kabeto, S.Y. Kim, and A.B. Rosen. 2008. "Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity?" *Alzheimer's & Dementia* 4(2):134-144.

Langa, K.M., B.L. Plassman, R.B. Wallace, A.R. Herzog, S.G. Heeringa, M.B. Ofstedal, J.R. Burke, G.G. Fisher, N.H. Fultz, M.D. Hurd, G.G. Potter, W.L. Rodgers, D.C. Steffens, D.R. Weir, and R.J. Willis. 2005. "The Aging, Demographics, and Memory Study: study design and methods." *Neuroepidemiology* 25(4):181-191.

Larson, E.B.and K.M. Langa. 2017. "What's the "Take Home" from Research on Dementia Trends?" *PLoS medicine* 14(3):e1002236.

Lau, B., S.R. Cole, and S.J. Gange. 2009. "Competing risk regression models for epidemiologic data." *Am J Epidemiol* 170(2):244-256.

Lawless, J.F.and N. Nazeri Rad. 2015. "Estimation and assessment of markov multistate models with intermittent observations on individuals." *Lifetime Data Anal* 21(2):160-179.

Leon, B.M.and T.M. Maddox. 2015. "Diabetes and cardiovascular disease: Epidemiology, biological mechanisms, treatment recommendations and future research." *World J Diabetes* 6(13):1246-1258.

Licher, S., S.K.L. Darweesh, F.J. Wolters, L. Fani, A. Heshmatollah, U. Mutlu, P.J. Koudstaal, J. Heeringa, M.J.G. Leening, M.K. Ikram, and M.A. Ikram. 2019. "Lifetime risk of common neurological diseases in the elderly population." *J Neurol Neurosurg Psychiatry* 90(2):148-156.

Lin, D.Y., E.J. Feuer, R. Etzioni, and Y. Wax. 1997. "Estimating medical costs from incomplete follow-up data." *Biometrics* 53(2):419-434.

Lin, K.A., K.R. Choudhury, B.G. Rathakrishnan, D.M. Marks, J.R. Petrella, P.M. Doraiswamy, and I. Alzheimer's Disease Neuroimaging. 2015. "Marked gender differences in progression of mild cognitive impairment over 8 years." *Alzheimer's & dementia (New York, N. Y.)* 1(2):103-110.

Lin, P.J., Y. Zhong, H.M. Fillit, E. Chen, and P.J. Neumann. 2016. "Medicare expenditures of individuals with Alzheimer's disease and related dementias or mild cognitive impairment before and after diagnosis." *Journal of the American Geriatrics Society* 64(8):1549-1557.

Livingston, G., A. Sommerlad, V. Orgeta, S.G. Costafreda, J. Huntley, D. Ames, C. Ballard, S. Banerjee, A. Burns, J. Cohen-Mansfield, C. Cooper, N. Fox, L.N. Gitlin, R. Howard, H.C. Kales, E.B. Larson, K. Ritchie, K. Rockwood, E.L. Sampson, Q. Samus, L.S. Schneider, G. Selbaek, L. Teri, and N. Mukadam. 2017. "Dementia prevention, intervention, and care." *Lancet* 390(10113):2673-2734.

Matthews, K.A., W. Xu, A.H. Gaglioti, J.B. Holt, J.B. Croft, D. Mack, and L.C. McGuire. 2019. "Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged >/=65 years." *Alzheimers Dement* 15(1):17-24.

Matthews, K.A., W. Xu, A.H. Gaglioti, J.B. Holt, J.B. Croft, D. Mack, and L.C. McGuire. 2019. "Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged≥ 65 years." *Alzheimer's & Dementia* 15(1):17-24.

Mayeda, E.R., M.M. Glymour, C.P. Quesenberry, and R.A. Whitmer. 2016. "Inequalities in dementia incidence between six racial and ethnic groups over 14 years." *Alzheimer's & Dementia* 12(3):216-224.

Mehta, K.M., K. Yaffe, E.J. Perez-Stable, A. Stewart, D. Barnes, B.F. Kurland, and B.L. Miller. 2008. "Race/ethnic differences in AD survival in US Alzheimer's Disease Centers." *Neurology* 70(14):1163-1170.

Mielke, M.M. 2018. "Sex and Gender Differences in Alzheimer's Disease Dementia." *The Psychiatric times* 35(11):14-17.

Naimi, A.I.and E.J. Tchetgen Tchetgen. 2015. "Invited commentary: Estimating population impact in the presence of competing events." *Am J Epidemiol* 181(8):571-574.

National Research Council. 2005. *Beyond the market: Designing nonmarket accounts for the United States*: National Academies Press.

Ng, T., C. Harrington, and M. Kitchener. 2010. "Medicare and medicaid in long-term care."

Health affairs (Project Hope) 29(1):22-28.

Norton, S., F.E. Matthews, D.E. Barnes, K. Yaffe, and C. Brayne. 2014. "Potential for primary prevention of Alzheimer's disease: an analysis of population-based data." *Lancet Neurol* 13(8):788-794.

Norton, S., F.E. Matthews, and C. Brayne. 2013. "A commentary on studies presenting projections of the future prevalence of dementia." *BMC Public Health* 13(1):1.

Ofstedal, M., D. Weir, K. Chen, and J. Wagner. 2011. "Updates to HRS sample weights (HRS Documentation Report DR-013)." *Ann Arbor: Survey Research Center, University of Michigan. Retrieved from//hrsonline. isr. umich. edu/sitedocs/userg/dr-013. pdf.*

Ofstedal, M., D. Weir, C. Kuang-Tsung, and J. Wagner. 2016. "Updates to HRS Sample Weights; 2011."

Ofstedal, M.B., G.G. Fisher, and A.R. Herzog. 2005. "Documentation of Cognitive Functioning Measures in the Health and Retirement Study."

Ortman, J.M., V.A. Velkoff, and H. Hogan. 2014. "An Aging Nation: The Older Population in the United States." *Current Population Reports, P25-1140. U.S. Census Bureau, Washington, DC. 2014.*

Oudshoorn, C.G.M. 1999. "Flexible multiple imputation by chained equations of the AVO-95 survey." *http://web. inter. nl. net/users/S. van. Buuren/mi/docs/rapport99045. pdf.*

Patterson, C. 2018. "World Alzheimer Report 2018: the state of the art of dementia research:

new frontiers." *Alzheimer's Disease International (ADI): London, UK.*

Pearl, J. 1995. "Causal diagrams for empirical research." *Biometrika* 82(4):669-688.

Pi-Sunyer, F.X., D.M. Becker, C. Bouchard, R. Carleton, G. Colditz, W. Dietz, J. Foreyt, R. Garrison, S.M. Grundy, and B. Hansen. 1998. "Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: Executive summary." *American Journal of Clinical Nutrition* 68(4):899-917.

Podcasy, J.L.and C.N. Epperson. 2016. "Considering sex and gender in Alzheimer disease and other dementias." *Dialogues Clin Neurosci* 18(4):437-446.

Pool, L.R., H. Ning, D.M. Lloyd-Jones, and N.B. Allen. 2017. "Trends in racial/ethnic disparities in cardiovascular health among US adults from 1999–2012." *Journal of the American Heart Association* 6(9):e006027.

Potter, G.G., B.L. Plassman, J.R. Burke, M.U. Kabeto, K.M. Langa, D.J. Llewellyn, M.A. Rogers, and D.C. Steffens. 2009. "Cognitive performance and informant reports in the diagnosis of cognitive impairment and dementia in African Americans and whites." *Alzheimers Dement* 5(6):445-453.

Prentice, R.L., J.D. Kalbfleisch, A.V. Peterson, Jr., N. Flournoy, V.T. Farewell, and N.E. Breslow. 1978. "The analysis of failure times in the presence of competing risks." *Biometrics* 34(4):541-554.

Prescott, B.T. 2019. "Recent Trends in Several Key Metrics in Higher Education." *Change: The*

Magazine of Higher Learning 51(1):22-27.

Prince, M., G.C. Ali, M. Guerchet, A.M. Prina, E. Albanese, and Y.T. Wu. 2016. "Recent global trends in the prevalence and incidence of dementia, and survival with dementia." *Alzheimers Res Ther* 8(1):23.

Prince, M., R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C.P. Ferri. 2013. "The global prevalence of dementia: a systematic review and metaanalysis." *Alzheimers Dement* 9(1):63-75 e62.

Prince, M.and J. Jackson. 2009. "World Alzheimer Report 2009." *London: Alzheimer's Disease International*:1-96.

Rajan, K.B., J. Weuve, L.L. Barnes, R.S. Wilson, and D.A. Evans. 2019. "Prevalence and incidence of clinically diagnosed Alzheimer's disease dementia from 1994 to 2012 in a population study." *Alzheimers Dement* 15(1):1-7.

Robins, J. 1986. "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect." *Mathematical modelling* 7(9-12):1393-1512.

Robins, J., M. Hernán, G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs. 2009. "Longitudinal data analysis." *by Fitzmaurice G., Davidian M., Verbeke G., Molenberghs G., Chapman and Hall/CRC Press, New York*:553-599.

Robins, J.M.and M.A. Hernán. 2009. "Estimation of the causal effects of time-varying

exposures." *Longitudinal data analysis* 553:599.

Robins, J.M., M.A. Hernán, and U. Siebert. 2004. "Effects of multiple interventions." *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors* 1:2191-2230.

Rocca, W.A., R.H. Cha, S.C. Waring, and E. Kokmen. 1998. "Incidence of dementia and Alzheimer's disease: a reanalysis of data from Rochester, Minnesota, 1975–1984." *American journal of epidemiology* 148(1):51-62.

Rocca, W.A., B.R. Grossardt, and L.T. Shuster. 2014. "Oophorectomy, estrogen, and dementia: a 2014 update." *Mol Cell Endocrinol* 389(1-2):7-12.

Roehr, S., A. Pabst, T. Luck, and S.G. Riedel-Heller. 2018. "Is dementia incidence declining in high-income countries? A systematic review and meta-analysis." *Clinical epidemiology* 10:1233.

Rogers, N.T., A. Steptoe, and D. Cadar. 2017. "Frailty is an independent predictor of incident dementia: Evidence from the English Longitudinal Study of Ageing." *Sci Rep* 7(1):15746.

Rusanen, M., M. Kivipelto, E. Levalahti, T. Laatikainen, J. Tuomilehto, H. Soininen, and T. Ngandu. 2014. "Heart diseases and long-term risk of dementia and Alzheimer's disease: a population-based CAIDE study." *J Alzheimers Dis* 42(1):183-191.

Ryan, C.L.and K. Bauman. 2016. "Educational attainment in the United States: 2015."

Salazar, J.C., F.A. Schmitt, L. Yu, M.M. Mendiondo, and R.J. Kryscio. 2007. "Shared random effects analysis of multi-state Markov models: application to a longitudinal study of transitions

to dementia." *Stat Med* 26(3):568-580.

Salib, E.and V. Hillier. 1997. "A case-control study of smoking and Alzheimer's disease." *International journal of geriatric psychiatry* 12(3):295-300.

Samieri, C., M.C. Perier, B. Gaye, C. Proust-Lima, C. Helmer, J.F. Dartigues, C. Berr, C. Tzourio, and J.P. Empana. 2018. "Association of Cardiovascular Health Level in Older Age With Cognitive Decline and Incident Dementia." *Jama* 320(7):657-664.

Samper-Ternent, R., Y.F. Kuo, L.A. Ray, K.J. Ottenbacher, K.S. Markides, and S. Al Snih. 2012. "Prevalence of health conditions and predictors of mortality in oldest old Mexican Americans and non-Hispanic whites." *J Am Med Dir Assoc* 13(3):254-259.

Satizabal, C.L., A.S. Beiser, V. Chouraki, G. Chene, C. Dufouil, and S. Seshadri. 2016. "Incidence of Dementia over Three Decades in the Framingham Heart Study." *N Engl J Med* 374(6):523-532.

Schaller, S., J. Mauskopf, C. Kriza, P. Wahlster, and P.L. Kolominsky-Rabas. 2015. "The main cost drivers in dementia: a systematic review." *Int J Geriatr Psychiatry* 30(2):111-129.

Schneider, J.A., Z. Arvanitakis, W. Bang, and D.A. Bennett. 2007. "Mixed brain pathologies account for most dementia cases in community-dwelling older persons." *Neurology* 69(24):2197-2204.

Selkoe, D.J. 2001. "Alzheimer's disease results from the cerebral accumulation and cytotoxicity of amyloid beta-protein." *J Alzheimers Dis* 3(1):75-80.

Sharp, E.S.and M. Gatz. 2011. "Relationship between education and dementia: an updated systematic review." *Alzheimer Dis Assoc Disord* 25(4):289-304.

Singh-Manoux, A., A. Dugravot, M. Shipley, E.J. Brunner, A. Elbaz, S. Sabia, and M. Kivimaki. 2018. "Obesity trajectories and risk of dementia: 28 years of follow-up in the Whitehall II Study." *Alzheimers Dement* 14(2):178-186.

Skoog, I. 2016. "Dementia: dementia incidence—the times, they are a-changing." *Nature Reviews Neurology* 12(6):316.

Sloane, P.D., S. Zimmerman, C. Suchindran, P. Reed, L. Wang, M. Boustani, and S. Sudha. 2002. "The public health impact of Alzheimer's disease, 2000–2050: potential implication of treatment advances." *Annual review of public health* 23(1):213-231.

Snyder, H.M., S. Asthana, L. Bain, R. Brinton, S. Craft, D.B. Dubal, M.A. Espeland, M. Gatz, M.M. Mielke, J. Raber, P.R. Rapp, K. Yaffe, and M.C. Carrillo. 2016. "Sex biology contributions to vulnerability to Alzheimer's disease: A think tank convened by the Women's Alzheimer's Research Initiative." *Alzheimers Dement* 12(11):1186-1196.

Sonnega, A., J.D. Faul, M.B. Ofstedal, K.M. Langa, J.W. Phillips, and D.R. Weir. 2014. "Cohort Profile: the Health and Retirement Study (HRS)." *International journal of epidemiology* 43(2):576-585.

Spirtes, P., C.N. Glymour, R. Scheines, D. Heckerman, C. Meek, G. Cooper, and T. Richardson. 2000. *Causation, prediction, and search*: MIT press.

Stern, Y. 2002. "What is cognitive reserve? Theory and research application of the reserve concept." *Journal of the international neuropsychological society* 8(3):448-460.

Stern, Y. 2009. "Cognitive reserve." *Neuropsychologia* 47(10):2015-2028.

—. 2012. "Cognitive reserve in ageing and Alzheimer's disease." *Lancet Neurol* 11(11):1006-1012.

Steyerberg, E.W., F.E. Harrell Jr, G.J. Borsboom, M. Eijkemans, Y. Vergouwe, and J.D.F. Habbema. 2001. "Internal validation of predictive models: efficiency of some procedures for logistic regression analysis." *Journal of clinical epidemiology* 54(8):774-781.

Suehs, B.T., S.N. Shah, C.D. Davis, J. Alvir, W.E. Faison, N.C. Patel, D. van Amerongen, and J. Bobula. 2014. "Household Members of Persons with A lzheimer's Disease: Health Conditions, Healthcare Resource Use, and Healthcare Costs." *Journal of the American Geriatrics Society* 62(3):435-441.

Taubman, S.L., J.M. Robins, M.A. Mittleman, and M.A. Hernan. 2009. "Intervening on risk factors for coronary heart disease: an application of the parametric g-formula." *International journal of epidemiology* 38(6):1599-1611.

Taylor Jr, D.H., T. Østbye, K.M. Langa, D. Weir, and B.L. Plassman. 2009. "The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited." *Journal of Alzheimer's disease: JAD* 17(4):807.

Taylor Jr, D.H.and F.A. Sloan. 2000. "How much do persons with Alzheimer's disease cost

Medicare?" *Journal of the American Geriatrics Society* 48(6):639-646.

U.S. Census Bureau. 2017. "Census 2017, National Population Projections, Main Series; generated by Jordan Weiss; using American FactFinder; <http://factfinder.census.gov>; (01 July 2019)."

—. 2018. "Census 2018, Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin; generated by Jordan Weiss; using American FactFinder; <http://factfinder.census.gov>; (01 July 2019)".

Van Buuren, S., J.P. Brand, C.G. Groothuis-Oudshoorn, and D.B. Rubin. 2006. "Fully conditional specification in multivariate imputation." *Journal of statistical computation and simulation* 76(12):1049-1064.

van der Flier, W.M.and P. Scheltens. 2005. "Epidemiology and risk factors of dementia." *J Neurol Neurosurg Psychiatry* 76 Suppl 5(suppl 5):v2-7.

Van Duijn, C.M.and A. Hofman. 1991. "Relation between nicotine intake and Alzheimer's disease." *Bmj* 302(6791):1491-1494.

Vangen-Lønne, A.M., P. Ueda, P. Gulayin, T. Wilsgaard, E.B. Mathiesen, and G. Danaei. 2018. "Hypothetical interventions to prevent stroke: an application of the parametric g-formula to a healthy middle-aged population." *European journal of epidemiology* 33(6):557-566.

Vuorinen, M., G. Spulber, S. Damangir, E. Niskanen, T. Ngandu, H. Soininen, M. Kivipelto, and A. Solomon. 2015. "Midlife CAIDE dementia risk score and dementia-related brain changes up

to 30 years later on magnetic resonance imaging." *J Alzheimers Dis* 44(1):93-101.

Weiner, M., N.R. Powe, W.E. Weller, T.J. Shaffer, and G.F. Anderson. 1998. "Alzheimer's disease under managed care: implications from Medicare utilization and expenditure patterns." *J Am Geriatr Soc* 46(6):762-770.

Weir, D.R. 2016. "Validating mortality ascertainment in the health and retirement study." *Ann Arbor, Michigan: Survey Research Center, Institute for Social Research, University of Michigan. Retrieved from https://hrs. isr. umich. edu/publications/biblio/9022*.

Westreich, D.and S.R. Cole. 2010. "Invited commentary: positivity in practice." *Am J Epidemiol* 171(6):674-677; discussion 678-681.

White, L., P. Fishman, A. Basu, P.K. Crane, E.B. Larson, and N.B. Coe. 2019. "Medicare expenditures attributable to dementia." *Health services research* 54(4):773-781.

Whitmer, R., D. Gustafson, E. Barrett-Connor, M. Haan, E. Gunderson, and K. Yaffe. 2008. "Central obesity and increased risk of dementia more than three decades later." *Neurology* 71(14):1057-1064.

Wilmoth, J.R., K. Andreev, D. Jdanov, D.A. Glei, C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon. 2007. "Methods protocol for the human mortality database." *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: http://mortality.org [version 31/05/2007]* 9:10-11.

Wilson, R.S., L.E. Hebert, P.A. Scherr, L.L. Barnes, C.F. Mendes de Leon, and D.A. Evans.

2009. "Educational attainment and cognitive decline in old age." *Neurology* 72(5):460-465.

Wortmann, M. 2012. "Dementia: a global health priority-highlights from an ADI and World Health Organization report." *Alzheimer's research & therapy* 4(5):40.

Wu, Y.-T., A.S. Beiser, M.M. Breteler, L. Fratiglioni, C. Helmer, H.C. Hendrie, H. Honda, M.A. Ikram, K.M. Langa, and A. Lobo. 2017. "The changing prevalence and incidence of dementia over time—current evidence." *Nature Reviews Neurology* 13(6):327.

Wu, Y.-T., A.S. Beiser, M.M.B. Breteler, L. Fratiglioni, C. Helmer, H.C. Hendrie, H. Honda, M.A. Ikram, K.M. Langa, A. Lobo, F.E. Matthews, T. Ohara, K. Pérès, C. Qiu, S. Seshadri, B.-M. Sjölund, I. Skoog, and C. Brayne. 2017. "The changing prevalence and incidence of dementia over time — current evidence." *Nature Reviews Neurology* 13:327.

Xu, W., C. Qiu, M. Gatz, N.L. Pedersen, B. Johansson, and L. Fratiglioni. 2009. "Mid- and late-life diabetes in relation to the risk of dementia: a population-based twin study." *Diabetes* 58(1):71-77.

Xu, W., L. Tan, H.F. Wang, M.S. Tan, L. Tan, J.Q. Li, Q.F. Zhao, and J.T. Yu. 2016. "Education and Risk of Dementia: Dose-Response Meta-Analysis of Prospective Cohort Studies." *Mol Neurobiol* 53(5):3113-3123.

Xu, W.L., A.R. Atti, M. Gatz, N.L. Pedersen, B. Johansson, and L. Fratiglioni. 2011. "Midlife overweight and obesity increase late-life dementia risk: a population-based twin study." *Neurology* 76(18):1568-1574.

Yaffe, K., C. Falvey, T.B. Harris, A. Newman, S. Satterfield, A. Koster, H. Ayonayon, E. Simonsick, and A.B.C.S. Health. 2013. "Effect of socioeconomic disparities on incidence of dementia among biracial older adults: prospective study." *Bmj* 347:f7051.

Yang, Z.and A. Levey. 2015. "Gender Differences: A Lifetime Analysis of the Economic Burden of Alzheimer's Disease." *Womens Health Issues* 25(5):436-440.

Yee, T.W. 2015. *Vector generalized linear and additive models: with an implementation in R*: Springer.

Young, J.G., L.E. Cain, J.M. Robins, E.J. O'Reilly, and M.A. Hernan. 2011. "Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula." *Statistics in biosciences* 3(1):119-143.

Zanchetti, A., C. Thomopoulos, and G. Parati. 2015. "Randomized controlled trials of blood pressure lowering in hypertension: a critical reappraisal." *Circ Res* 116(6):1058-1073.

Zhong, G., Y. Wang, Y. Zhang, J.J. Guo, and Y. Zhao. 2015. "Smoking is associated with an increased risk of dementia: a meta-analysis of prospective cohort studies with investigation of potential effect modifiers." *PloS one* 10(3).

Zhu, C.W., S. Cosentino, K. Ornstein, Y. Gu, N. Scarmeas, H. Andrews, and Y. Stern. 2015. "Medicare Utilization and Expenditures Around Incident Dementia in a Multiethnic Cohort." *J Gerontol A Biol Sci Med Sci* 70(11):1448-1453.

Zhu, C.W., N. Scarmeas, K. Ornstein, M. Albert, J. Brandt, D. Blacker, M. Sano, and Y. Stern.

2015. "Health-care use and cost in dementia caregivers: Longitudinal results from the Predictors Caregiver Study." *Alzheimer's & Dementia* 11(4):444-454.

Zhu, C.W., N. Scarmeas, R. Torgan, M. Albert, J. Brandt, D. Blacker, M. Sano, and Y. Stern. 2006. "Clinical characteristics and longitudinal changes of informal cost of Alzheimer's disease in the community." *J Am Geriatr Soc* 54(10):1596-1602.

Zhu, C.W., R. Torgan, N. Scarmeas, M. Albert, J. Brandt, D. Blacker, M. Sano, and Y. Stern. 2008. "Home health and informal care utilization and costs over time in Alzheimer's disease." *Home Health Care Serv Q* 27(1):1-20.

Zissimopoulos, J., E. Crimmins, and P. St Clair. 2014. "The Value of Delaying Alzheimer's Disease Onset." *Forum Health Econ Policy* 18(1):25-39.

Zissimopoulos, J.M., B.C. Tysinger, P.A. St Clair, and E.M. Crimmins. 2018. "The Impact of Changes in Population Health and Mortality on Future Prevalence of Alzheimer's Disease and Other Dementias in the United States." *J Gerontol B Psychol Sci Soc Sci* 73(suppl_1):S38-S47.