Dementia research under a causal inference lens

Liliana Paloma Rojas Saunero

2021-11-09

2

Contents

Chapter 1

Introduction

With the extension of life expectancy over the last few decades, dementia has become a major burden that affects the elderly. In 2016, the global number of individuals who lived with dementia was over 40 million(**gbd2016**) and by 2050 this number is expected to triple(**worldreport2018**). In 2020, deaths due to Alzheimer's disease and other dementias have increased compared to previous years, becoming the 7th leading cause of death globally and overtaking stroke to become the second leading cause in high-income countries. Although dementia risk does not represent a leading cause of disease and death in low-middle income countries, it is projected to increase as the burden of preventable and curable diseases reduce(**who2020**). Likewise, women are disproportionately affected, worldwide(**women2015**). Furthermore, the burden of dementia does not only affect those who have the disease: dementia has large effects on the lives of caregivers, families, and health-care systems. Unequal distribution of opportunities, responsibilities and societal roles push women into the caregiver role more than men, which hampers even more their access to paid work and health, especially in poor and marginalized areas, creating negative feedback loops that increase all kinds of disparities(**swinkels2019**; **brodaty2009**; **etters2008**). However, these harmful consequences can be prevented if the burden of dementia is reduced, and this can be done by targeting prevention and delay of onset(**carrillo2013**).

The challenge to reduce the burden of dementia is that this disease is complex and heterogeneous, with multiple etiologic and neuropathological processes likely related. The field is constantly focused on identifying the modifiable risk factors and therapies. To this purpose, in 2020, the Lancet Commission released an updated guideline with evidence on twelve modifiable risks factors, which would account for 40% of worldwide dementias that could have been prevented or delayed. These modifiable risk factors include: lower education, hypertension, hearing impairment, smoking, obesity, depression, physical inactivity, diabetes, low social contact, alcohol consumption, traumatic brain injury (TBI), and air

pollution(**lancet2020**). Furthermore, there is a growth of brain biomarker research, with the simultaneous intention of detecting proxies for early diagnosis, as well as identifying new molecular targets for intervention. Since there are few specific drugs that target amyloid and tau production, with small-to-no evidence of effect, new studies for drugs that target other mechanisms are in study. Drug repositioning and repurposing research is becoming more popular, offering a valuable alternative route for the identification of effective disease-modifying treatments for dementia(**ballard2020**; **langedijk2015**). In 2020, Cummings et al. identified 121 agents in clinical trials for the treatment of Alzheimer's disease, out of which 43% (57) represented repurposed agents across all phases of the pipeline(**cummings2020**).

All these objectives in dementia research are heavily reliant on observational studies and current availability of multiple sources of large data give us the opportunity to expand the field. Nevertheless, new data sources and sophisticated computational software and technology, may sometimes overshadow the process of asking clear questions and the steps to tie the questions to the data. In a time where we have deeply embraced that "causation is not correlation", the concept of "associations" and molding questions to hypothesis testing has deeply overshadowed the critical step in research to clearly define questions and wanted interpretations. Acknowledging that etiologic research in dementia research is aimed at identifying causal effects is probably an underlying challenge and impediment to conceptualize clear causal questions. This may be due to the misconception that causality can only be inferred from randomized controlled studies (RCTs), which is reinforced by high-impact journals(**jama**). Another reason why researchers may struggle to embrace causal thinking is due to the polarized debates among causal inference experts about how to conceptualize exposures that cannot be intervened upon or manipulated (such as sex, race or BMI)(**waterkills**; **schwartz2016**). During my first years of training in causal inference, this debate confused me and filled me with insecurities about how to study exposures related to dementia when I did not have measurements on interventional data. Nevertheless, the reader may find in the following chapters how I embraced this debate and developed my own criteria on how to study causal questions.

In this thesis, I aimed to study the effect of several potential targets of intervention related to dementia prevention that have had controversial results in previous observational studies, by applying causal inference theory and corresponding methods. Over the following chapters I will describe how the creative process of thinking and conceptualizing what we truly want to ask helps to formulate clear questions, identify potential sources of bias, and articulate the analytic methods of available data to match the question. To this matter, will implement the target trial emulation framework(**hernan_robins_2016**; **labrecque2018**) and, to answer each question, I will use data collected for the Rotterdam Study, a population-based cohort study with rich longitudinal data assessments(**ikram2020**), described in the next section.

To give further context, in the decade of the sixties, Cochran reinforced the notion that observational studies were aimed to answer questions about cause-effect when randomized studies were not feasible or ethical(**cochran1965**). In the eighties, Robins and others[@] formalized how to answer causal questions with observational studies following the RCT principles, and decades after the "target trial emulation framework" was branded(**hernan_robins_2016**). Labrecque and Swanson define the target trial emulation as "the application of design principles from randomized trials to the analysis of observational data, explicitly tying the analysis to the trial it is emulating"(**labrecque2018**). This process includes specifying and emulating the key components of the trial protocol such as the eligibility criteria, treatment strategies, treatment assignment, follow-up period, outcome, causal contrasts and statistical analysis. The refinement of the causal question, and the target trial specifications will often be a back and forward process between the question we truly aim to answer and the availability of data to answer that question(**labrecque2018**; **whatif2020**).

Taking in consideration this framework and principles, in **Chapter ??** I present an emulation of a target trial to study the effect of statins treatment in the 10-year risk of dementia and death. This work brings clarity to the idea that even in observational studies we can formulate causal contrasts like the intention-to-treat and the per-protocol effect, as in pragmatic trials. In this setting, the intention-to-treat effect refers to a combination of the effect of initiating the treatment under study and of any other patient and physician's behavioural changes triggered by the assignment itself. This effect is agnostic of any treatment decisions made after baseline, which makes it difficult to interpret to patients, clinicians and other decision-makers(**murray2019**; **murray2018**). Thought this is what makes it appealing from an analytic perspective, since it can be conceptualized as a point treatment strategy. Instead, the per-protocol effect represents the effect of being assigned a treatment strategy and adhering to that assigned treatment strategy through-out follow-up, as specified in the study protocol. This effect can be conceptualized as a static or dynamic treatment strategy, since adherence to a treatment strategy over follow-up will depend on the evolution of an individual's time-varying covariates(**whatif2020**). Being explicit about the treatment strategy emphasizes the necessity to collect data on the treatment adherence over follow-up, as well as time-varying confounders and predictors of adherence. It also introduces the major challenge with time-varying treatments, the treatment-confounder feedback.

Treatment-confounder feedback refers to the setting where time-varying covariates affect treatment over time, but additionally, time-varying covariates are affected themselves by prior treatment(**robins1986**; **whatif2020**). If we could conceptualize the modifiable risks factors proposed by the Lancet commission as potential targets of intervention, and define interventions as static or dynamic, we must conceptualize the potential treatment-confounder feedback loops that might challenge both randomize trials and observational studies. For example, the Lancet commission defines hypertension as one of the modifiable risk factors. In **Chapter ??** I illustrate how to conceive a question where the in-

terest is focused in learning how much would the risk of stroke and dementia change if, hypothetically, we could reduce and keep systolic blood pressure under different thresholds defined in clinical practice over follow-up, resembling a per-protocol effect. Given that blood pressure and hypertension is affected by other comorbidities, and it affects other comorbidities as well, proper analysis that accounts for treatment-confounder feedback was needed. Since traditional statistical methods cannot account for this feature, one of the highlights of this dissertation is the application of "*G-methods*". G-methods (or generalized methods) are a set of causal models and analytic methods proposed by Robins beginning in 1986, consisting in the g-formula, marginal structural models and structural nested models(**robins1986**; **whatif2020**). These methods have revolutionized the field of epidemiology and public health, by providing analytic tools to answer causal questions with longitudinal treatment/exposure data(**richardson2014**).

In addition to the target trial framework and the G-methods, another key causal inference tool is the development and application of causal directed acyclic graphs (DAGs). These causal diagrams are another representation of causal structures which allows to draw and visualize assumptions. DAGs were introduced by Judea Pearl(**pearl1994**) and extended by Robins to settings of time-varying exposures(**robins1986**). In **Chapter ??chapter4)** I progressively build a DAG to represent a causal question of the effect of a Pin1-targeting drug on the risk of dementia, when we only have cancer diagnosis measured as the proxy for Pin-1. This project was motivated to understand all potential sources of bias that could be related to inverse association between cancer diagnosis and dementia that had been systematically reported in previous observational studies. Since we cannot understand bias if we do not have a clear causal question to start with, causal graphs helped elucidate all the steps to connect the unmeasured mechanism of interest to the observed data outlining the data generation process.

Just as it is important to clearly define what we mean by the exposure or intervention of interest, we must focus on other elements of a clear research question, such as how our question incorporates competing events and other censoring events. Given that more than 30% of the participants died prior to dementia diagnosis in the Rotterdam Study, death played a major role as a competing event in all the projects of this dissertation. At the same time I was drafting one of the first projects of this dissertation, a pre-print about competing events in a causal inference framework, by Young et al.(**young2020**) was published. This paper helped me understand that death was not something to "fix" within the analysis of data, but rather to include as part of the question. This goes in hand with the definition of "estimands" by the ICH9 addendum(**ich9**) that considers post-randomization events, defined as intercurrent events, that can affect outcome assessment or interpretation, as part of this research question.

As opposed to previous recommendations that suggest to use a "cause-specific hazard model" for etiologic research, and a "Fine and Gray sub-distribution

hazard model" for prediction research(**lau2009**), Young et al.(**young2020**) disengage from this recommendation, and rather start by framing the different estimands that are approached by traditional methods in survival analysis. Young et al. formalized how, under explicit assumptions, they allow for identification of different estimands, and present both directed acyclic graphs and single intervention world graphs for settings with competing events. The two causal questions or estimands discussed are the "controlled direct effect" and the "total effect". The controlled direct effect corresponds to a question where death could have been fully prevented, hypothetically. The total effect corresponds to a question where death can also happen through-out the follow-up, thus it captures the effect mediated through death. In each chapter of this dissertation, I apply either of these estimands (or both), depending on the research aim, demonstrating that different estimands will be better suited for different context. In **Chapter ??** I present a systematic review to describe the current practices and interpretations of longitudinal studies of dementia, where death plays the role of a competing event. Unfortunately, this work highlights how much in real applications censoring is treated as a synonym for ignoring and the large gap between methods development and applied research. This reality motivated me to write **Chapter ??**, were I will highlight the key conceptual definitions in regards to competing events in causal inference, with an applied example on smoking cessation in the risk of dementia, to make this novel framework accessible to applied researchers in the field.

To finalize, in **Chapter ??** I distill all the learning experiences from overcoming different methodological challenges in each project, the broader implications of my research and discuss the unsolved challenges as future lines of research. And, as a way to express my growth as a researcher and epidemiologist within this PhD trajectory, in **Chapter ??** I will zoom out from the current work in dementia research and discuss more broadly how the SARS-CoV-2 pandemic has shaped my understanding of becoming an epidemiologist interested in methods development.

1.1   Study setting

All the projects in this dissertation were designed and implemented using data from The Rotterdam Study, a population-based cohort that recruited participants living in the district of Ommoord, in Rotterdam, the Netherlands. The cohort was recruited from all inhabitants aged 55 years and older. Participants were invited in random clusters, through sampling from the municipal register(**hofman1991**). The Rotterdam Study, also known in the Netherlands as Erasmus Rotterdam Gezondheid Onderzoek (ERGO), started as a pilot study in July, 1989 and full recruitment and complete data acquisition started in January, 1990 and finished in September, 1993. Out of 11850 eligible participants, 7983 accepted (78% response rate). Participants were followed every three years, follow-up visits were held between 1993-1995, 1997-1999, 2002-2005, 2008-2014; however, time between visits varied for each participant ranging from 1 to 7 years between them.

In 1999, members of Ommoord who had become 55 years of age or moved into the study district since the recruitment were invited to participate; out of 4472 invitees, 3011 new participants were included to the Rotterdam Study(**hofman2007**) between February 1999 and December 2001, with follow-up visits between 2002-2005 and 2008-2012. In 2006, the cohort expanded to with a new wave of recruitment, including Ommord participants who were 40 years and older. This recruitment represented 3932 new participants in the Rotterdam Study. Last, between 2016 and December 2020, a new wave of recruitment included 3005 new participants. For this reason, the total sample size of the Rotterdam Study is around 18000 participants, though this number does not represent the total sample size observed at a specific time-point, due to the active recruitment and since several participants may have died or were loss-to follow-up before the latter waves of recruitment. Data collection dates and representation of the recruitment waves included in this dissertation with corresponding sample size are presented in Figure 1.

When the Rotterdam Study was conceived, it was focused on four primary areas of research: neurogeriatric diseases, cardiovascular locomotor and ophthalmologic diseases, though it later expanded to explore several other areas of disease. Hofman et al defined the aims of the study as follows: 1) to investigate the determinants of diseases in order to assess their etiologic significance; 2) to investigate potentially modifiable determinants in order to be able to develop preventive strategies by providing specific recommendations for intervention studies(**hofman1991**). Thus, participants went through extensive examinations, including interviews, physical examination and collecting bio-samples for molecular and genetic studies. Furthermore, the Rotterdam Study integrated information from secondary sources; for example, data on death was collected from municipal registries from the early start. Since 1991, data from all pharmacies serving in Ommoord region was integrated and, linkage to the Dutch Hospital Data, which captures main discharge diagnosis from all nationwide

hospital admissions was stablished in 1998. Several other examples of data integration are discussed by Ikram et. al(**ikram2020**). The immense effort to capture all this information about participants makes the Rotterdam Study a unique source to answer questions about time-varying exposures or interventions that could reduce the risk of dementia and other age-related diseases.

1.2 References

Chapter 2

Emulating a target trial of statin use and risk of dementia using cohort data

2.1   Abstract

Objective: Observational data can be used to attempt to emulate a target trial of statin use and estimate analogues of intention-to-treat and per-protocol effects on dementia risk.

Methods: Using data from a prospective cohort study in the Netherlands, we conceptualized a sequence of "trials" in which eligible individuals ages 55-80 years were classified as statin initiators or non-initiators for every consecutive month between 1993 and 2007 and were followed until diagnosis of dementia, death, loss to follow-up, or the end of follow-up. We estimated two types of effects of statin use on dementia and a combined endpoint of dementia or death: the effect of initiation versus no initiation and the effect of sustained use versus no use. We estimated risk by statin treatment strategy over time via pooled logistic regression. We used inverse-probability weighting to account for treatment-confounder feedback in estimation of per-protocol effects.

Results: Of 233,526 eligible person-trials (6,373 individuals), there were 622 initiators and 232,904 non-initiators. Comparing statin initiation with no initiation, the 10-year risk differences (95% CI) were -0.1% (-2.3%, 1.8%) for dementia and 0.3% (-2.7%, 3.3%) for dementia or death. Comparing sustained statin use versus no use, the 10-year risk differences were -2.2% (-5.2%, 1.6%) for dementia and -5.1% (-10.5%, -1.1%) for dementia or death.

Conclusions: Individuals with sustained statin use, but not statin initiation alone, had reduced 10-year risks of dementia and dementia or death. Our results should be interpreted with caution due to the small number of initiators and events, and potential for residual confounding.

2.2 Introduction

The effect of a commonly prescribed medication on reducing the long-term risk of chronic diseases is often of interest in public health research. However, estimating the effect of a medication for primary prevention can be challenging because it typically requires enrolling disease-free asymptomatic adults and following them for several years. Moreover, the effect of sustained medication use may be of great value for informing personal, clinical, or public health decision-making, even though randomized clinical trials often emphasize or solely estimate an intention-to-treat effect regardless of sustained use(**hernan__itt2012**).

For example, use of statin therapy in adults without cognitive impairment could change dementia risk later in life. The effects of statin use for primary and secondary prevention of cardiovascular disease in a general population are well established(**naci2013**; **taylor2013**; **udell2006**). However, randomized trials(**mrc2002**; **trompet2010**) and several observational studies assessing the association between statin use and Alzheimer's disease or dementia have had conflicting results ranging from statins potentially increasing, decreasing, or having a negligible effect on these outcomes. Overall, the randomized trials have included individuals at high risk of vascular disease, have had relatively short follow-up times (5 years or less), and have reported the intention-to-treat effect only(**mrc2002**; **trompet2010**). Meanwhile, comparable observational studies have notable limitations(**power2015**). For example, studies that only assess statin use at baseline are by definition limited by the number of individuals who are statin users at baseline(**arvanitakis2008**; **szwast2007**; **zandi2005**; **ancelin2012**; **smeeth2009**; **wolozin2007**), studies that assess statin use at the time of incident dementia (or with a 1-year or 2-year lag) could be particularly susceptible to reverse causation(**haag2009**; **cramer2008**; **bettermann2012**; **beydoun2011**; **li2004**; **li2010**; **rea2005**), studies that assess statin use as a time-varying variable using stratification-based techniques could be susceptible to bias due to inappropriate adjustment for time-varying confounding(**bettermann2012**; **beydoun2011**; **bernick2005**; **sparks2008**; **steenland2013**; **starr2004**; **cox2010**), and studies that include prevalent statin users could be susceptible to selection bias(**arvanitakis2008**; **szwast2007**; **zandi2005**; **ancelin2012**; **wolozin2007**; **haag2009**; **cramer2008**; **beydoun2011**; **li2004**; **li2010**; **rea2005**; **bernick2005**; **steenland2013**; **starr2004**). Finally, methods to estimate the effect of sustained statin treatment over follow-up time in addition to the effect of initiating statin treatment are underutilized in analyses of randomized trials and observational studies alike(**ray2003**).

To overcome some of the challenges of randomized clinical trials, we emulate a hypothetical randomized trial – a target trial – for estimating observational analogues to intention-to-treat and per-protocol effects of statin on dementia[28]. Through explicitly describing and emulating the target trial, we can leverage the richness of observational data while avoiding selection and residual confounding

biases that are consequences of common flaws in observational studies' design and analyses. The advantages of the target trial framework may be particularly pertinent in pharmacoepidemiologic research where time-varying variables can be both causes and consequences of treatment. Here, we describe the protocol of the target trial and then how to emulate it using observational data from the Rotterdam Study. Since few individuals initiate statins in a given calendar month, we emulate a sequence of target trials where, at each calendar month, eligibility criteria are applied anew and eligible individuals are then assigned to statin initiation or non-initiation (**danaei2013**; **danaei2018**).

2.3   Methods

Standard Protocol Approvals, Registrations, and Patient Consents

The medical ethics committee of the Erasmus Medical Centre approved the study. Written informed consent was obtained from all patients participating in the study.

Study population

This study is embedded in the Rotterdam Study, a prospective cohort study initially including 7983 individuals aged 55 years or older living in Ommoord, a district of Rotterdam, the Netherlands(**hofman2015**). Individuals living in the district were invited to participate in the cohort between 1990 and 1993. Home visits and center visits were conducted at enrollment (1990-1993) and again at follow-up visits in 1993-1995, 1997-1999, 2002-2004, 2009-2011, and 2014-2015. Demographic, clinical, and lifestyle factors were measured and recorded at each visit. In 2000, 3011 individuals who had become 55 years of age or moved into the study district since the start of the study were added to the cohort. Home visits and examinations were conducted at enrollment (2000-2001) and again at follow-up visits in 2004-2005, 2011-2012, and 2014-2015.

Statin measurement

Complete information on all prescriptions filled at any of seven automated pharmacies serving the Ommoord area (>99% of participants) were available for individuals enrolled in the Rotterdam Study starting on January 1, 1991(**hofman2015**). Simvastatin, pravastatin, fluvastatin, atorvastatin, cerivastatin, and rosuvastatin were classified as statins based on Anatomical Therapeutic Chemical (ATC) codes. Statins were approved in 1990 onwards in the Netherlands, with simvastatin the most commonly prescribed statin(**mantel2002**). The duration of a prescription was calculated as the total number of delivered units divided by the prescribed daily number of units. The date of delivery and duration of prescription were then used to calculate the number of treated days during each month for each individual. In the Netherlands, all statins are available only by prescription.

Dementia measurement

Participants were screened for dementia at baseline and subsequent center visits with the Mini-Mental State Examination (MMSE) and the Geriatric Mental Schedule organic level(**debruijn2015**). Those with an MMSE score <26 or Geriatric Mental Schedule score >0 underwent further investigation and informant interview, including the Cambridge Examination for Mental Disorders of the Elderly. All participants also underwent routine cognitive assessment at each center visit. In addition, the entire cohort was continuously under surveillance for clinically-diagnosed dementia through electronic linkage of the study database with medical records from general practitioners and the regional institute for outpatient mental health care. Available information on clinical neuroimaging was used when required for diagnosis of dementia subtype. A consensus panel led by a consultant neurologist established the final diagnosis and subtype of dementia according to standard criteria for dementia (DSM-III-R), Alzheimer's disease (NINCDS–ADRDA), and vascular dementia (NINDS-AIREN).

2.3.1 The protocol of the target trial

We begin by describing the protocol of our target trial to estimate the effect of statins on incident dementia; subsequently, we describe how we emulate this target trial in our data. A summary of these protocol components can be found in Table 1.

Eligibility criteria

The target trial includes participants aged 55-80 years with no statin prescription in the previous two years and no previous diagnosis of dementia. Individuals receive a cholesterol test and MMSE examination at enrollment, and are excluded if they have an MMSE score <26. Individuals are enrolled in the target trial starting in January 1993.

Treatment strategies

Two treatment strategies are considered: (1) initiate statin therapy at baseline and remain on statins during the follow-up; or (2) refrain from taking statin therapy during the follow-up. Both strategies allow for deviation if a serious illness occurred, e.g., cancer or heart disease.

Randomized assignment

Eligible individuals are randomized to one of the two treatment strategies without blinding.

Outcomes

The primary outcomes of interest include incident dementia and a composite outcome of incident dementia or all-cause mortality. Mortality is ascertained

through linkage with records of general practitioners and municipality records. Incident dementia is ascertained by combining continuous surveillance through electronic linkage with medical records and MMSE and Geriatric Mental Schedule assessments, using the algorithm described above.

Follow-up period

Individuals are followed from baseline (randomization) until diagnosis of dementia, mortality, loss to follow-up (defined by not attending regular study visits, e.g., 2 years without a study visit), or January 1st, 2015, whichever occurs earliest.

Causal contrasts of interest

To compare the two treatment strategies, we estimate a modified intention-to-treat effect as well as the per-protocol effect. In the target trial, the modified intention-to-treat effect is the effect of statin initiation versus no statin initiation (the intention-to-treat effect would be the effect of assignment to one of these strategies). The per-protocol effect is the effect of statin initiation and sustained statin use versus no statin initiation and never initiating statin medications.

Analysis plan

Modified intention-to-treat effect: For each outcome, we fit the pooled logistic regression model: $logit Pr(D_{t+1} = 1|D_t = 0, A) = \theta_{0t} + \theta_1 A + \theta_2 A * h(t)$ where D_t is an indicator for developing the outcome during time t (1: yes, 0: no), A is an indicator for treatment initiation (1: statins; 0:no statins), θ_{0t} is a time-varying intercept and $h(t)$ for follow-up time, modeled as restricted cubic splines. The model's predicted values are then used to estimate 10-year dementia-free and 10-year dementia-free survival curves, and 5- and 10-year risks of dementia and the combined endpoint of dementia and death.

Per protocol effect: Individuals are artifically censored when they deviate from their assigned treatment strategy. Specifically, individuals assigned to initiation and sustained statin use are censored after one calendar month with no treated days, and individuals assigned to refrain from statins are censored after one calendar month with one or more treated days. Individuals can no longer be artificially censored after a diagnosis of heart disease or cancer.

For each outcome we fit a weighted pooled logistic regression model: $logit Pr(D_{t+1} = 1|D_t = 0, C_t = 0, A, V) = \theta_{0t} + \theta_1 A + \theta_2 A * h(t) + \theta_3' V$ where C_t is an indicator for artificial censoring at time t (1: yes, 0: no) and V is a vector of covariates measured at baseline. To adjust for time-varying selection bias induced by the artificial censoring, we weight each individual at each time t by the inverse probability of receiving their own time-varying history [34, 35]: i.e., the weight

$$W_t = \prod_{k=0}^{t} \frac{1}{f(A_k \mid \bar{A}_{k-1}, \bar{D}_k = 0, \bar{L}_{k-1})}$$

where $f(A_k|\bar{A}_{k-1}, \bar{D}_k = 0, \bar{L}_{k-1})$ is the conditional probability mass function $f(A_k|\bar{A}_{k-1}, \bar{D}_k = 0, \bar{L}_{k-1})(a_{k-1}|\bar{a}_{k-1}, d_k = 0, \bar{l}_{k-1})$ with $(a_{k-1}|\bar{a}_{k-1}, a_k = 0, \bar{l}_{k-1})$ evaluated at the random argument $(A_k|\bar{A}_{k-1}, \bar{D}_k = 0, \bar{L}_{k-1})$. We use overbars. We use overbars to denote an individual's covariate or treatment history. For an individual who is uncensored through time t, note that the contribution to the denominator of the weight is equal to the probability that the individual remains uncensored through time t conditional on not developing the outcome by time t, covariate history, and treatment history [36]. An individual's weight at each time is therefore the cumulative product of the conditional probability of remaining uncensored. The weights are estimated by fitting a pooled logistic model including a time-specific intercept, the baseline covariates and the most recent measurement of several time-varying covariates.

Baseline covariates included variables measured at the clinic and home visits: sex, highest educational attainment, calendar year, age, APOE- 4 carrier status, body mass index (BMI), most recent MMSE, most recent total cholesterol measurement, most recent systolic blood pressure measurement, and current smoking status; variables measured using clinical diagnoses: history of stroke, history of heart disease, and history of cancer; variables measured using continuous pharmacy dispensing records: antihypertensives, diuretics, beta-blockers, calcium blockers, renin angiotensin aldosterone system (RAAS) inhibitors, nonstatin cholesterol-lowering medications, NSAIDs, psychotropic medications (antidepressants, antipsychotics, and benzodiazepines), and aspirin (yes, no); and variables measured using a combination of pharmacy records and clinical diagnoses: history of diabetes. Time-varying covariates that could be common causes of statin initiation (or discontinuation) and incident dementia included MMSE score, months since most recent MMSE, total cholesterol, months since the most recent cholesterol measurement, systolic blood pressure, current smoking status, history of stroke, heart disease, cancer, and diabetes, and use of other prescription medications. All continuous variables were modeled linearly. The contributions to the weights are set to 1 after diagnosis of heart disease or cancer. Weights are truncated at the 99th percentile[37] (~7.8, Appendix Table 3, available from Dryad doi:10.5061/dryad.kprr4xh19) .

Subgroup analyses

We explore effect modification by calendar year, cohort, and age in subgroup analyses restricted to: (1) calendar years in 2000 or later; (2) the initial cohort only; (3) individuals younger than 70 years of age at baseline; and (4) individuals 70 years of age or older at baseline.

Competing events

In our setting, death is a competing event for dementia because an individual cannot get dementia once they have died. The definition and interpretation of the causal estimand for an effect of statins on dementia is tied to the analytic choice to define the competing event death as a censoring event (i.e. an event that prevents observation of the counterfactual outcome of interest) or not[38]. In

our dementia analyses, we consider the competing event death to be a censoring event. That is, our modified intention-to-treat and per-protocol estimates are interpreted as the effect of statin initiation and sustained statin use on incident dementia in a hypothetical population in which death does not occur or in which death is independent of risk factors for dementia (conditional on measured covariates). This estimate can also be interpreted as the controlled direct effect of statins on dementia not mediated by the competing event death [39]. To adjust for potential bias due to competing events in the per-protocol analysis, we compute inverse probability weights for death in a sensitivity analysis. Each individual receives a timevarying weight inversely proportional to the probability of not dying. To compute the weights, we fit a pooled logistic model including a time-specific intercept, the baseline covariates and time-varying covariates listed previously, and an indicator for the treatment arm. The product of the estimated conditional probabilities at each time is then used to estimate the time-varying weight for each person at each time. Another approach for dealing with competing risks is to consider a composite outcome of dementia or death. Although this approach addresses some issues with the interpretation and valid estimation of our primary approach, it remains difficult to interpret because a non-null estimate can occur due to an effect of statins on dementia alone, death alone, or a combination thereof.

Loss to follow-up

To adjust for potential selection bias due to non-differential loss to follow-up between the treatment arms, we compute inverse probability of censoring weights in a sensitivity analysis. Each individual receives a time-varying weight inversely proportional to the probability of not being loss to follow-up.

2.3.2 Emulation of the target trial using observational data

We emulated the target trial by using Rotterdam Study data from the initial and first extended cohort. We will first describe emulating a target trial with baseline in January 1993, and then describe how we can emulate a series of target trials with baseline months ranging from January 1993 to December 2007.

Eligibility criteria

We identified individuals aged 55-80 years with no statin prescription (operationalized as at least 2 years enrolled in the Rotterdam Study with no recorded statin prescription) and no previous diagnosis of dementia. Individuals were excluded if they did not have a recent cholesterol measurement (within the previous 3 years) if they did not have a recent MMSE score 26 (within the previous 3 years), or if they did not have any BMI or SBP measurement.

Treatment strategies

Eligible individuals were classified as statin initiators if they initiated statins in January 1993 and as non-initiators if they did not initiate statins in January 1993.

Randomized assignment

To emulate the randomization component of the target trial, we adjusted for the baseline variables listed in our analysis (see below).

Follow-up and Outcome

Same as in the target trial, except we define loss to follow-up as ten years without an MMSE measurement (due to infrequent scheduled visits in the later years of the Rotterdam Study).

Causal contrasts

Same as in the target trial.

Analysis plan

The modified intention-to-treat analysis was the same as in the target trial except that we included the previously described baseline covariates in the pooled logistic regression model. Because individuals in the observational study did not necessarily have MMSE and cholesterol measurements taken exactly at baseline, we also included time since most recent MMSE measurement and time since most recent cholesterol measurement as baseline covariates. The per-protocol analysis was the same as in the target trial (individuals are artificially censored when they deviate from their assigned treatment strategy) except that we included the two additional baseline covariates time since most recent MMSE measurement and time since most recent cholesterol measurement as baseline covariates.

Sensitivity analyses and subgroup analyses

We performed the same sensitivity and subgroup analyses as in the target trial. Further, we conducted several additional sensitivity analyses. To evaluate residual confounding by indication, we excluded individuals who: (1) had a history of heart disease or history of stroke at baseline; (2) were taking vascular medications at baseline (non-statin cholesterol-lowering medications, diuretics, beta-blockers, calcium-blockers, RAAS inhibitors, NSAIDs, or aspirin); (3) had high cholesterol (>6.2 mmol/L) at baseline; and (4) did not have high cholesterol at baseline. To evaluate whether our results were sensitive to overly strict inclusion criteria that could decrease our sample size, we required a 6-month statin washout period at baseline (rather than two-year) and required an MMSE 26 at baseline but did not require it to be in the previous 3 years.

Creating a sequence of target trials using observational data

Enrollment in the target trial begins in January 1993 and continues until the desired sample size has been attained. In the emulation of the target trial described above, the eligibility criteria are applied once, in January 1993. However,

if only a small number of eligible individuals initiate statin therapy in January 1993, a meaningful comparison of statin initiators and non-initiators will not be possible. To increase the sample size in our trial emulation, we apply the same eligibility criteria anew in February 1993 and each month thereafter until December 2007. That is, we emulate a series of 180 'trials', each of them with a 1-month enrollment period[30]. For example, individuals eligible to enroll in the January 1993 'trial' who do not initiate statins that month may be eligible to enroll in the February 1993 'trial', assuming they continue to meet the other eligibility criteria that month.

In the analysis, baseline variables are updated at the start of each 'trial'. For example, the baseline variables for the February 1993 'trial' are the most recent measurements of the covariates at that time. We pooled data from all 180 'trials' into a single model and include 'month at the trial's baseline' (taking values from 1 to 180) and month of follow-up in each 'trial' in our models (both modeled using restricted cubic splines).

All analysis where performed in SAS using publicly available macros (https://www.hsph.harvard.edu/causal/software/). 95% confidence intervals were obtained using bootstrapping with 200 samples.

Data Availability

Data can be obtained on request. Requests should be directed toward the management team of the Rotterdam Study (secretariat.epi@erasmusmc.nl), which has a protocol for approving data requests. Because of restrictions based on privacy regulations and informed consent of the participants, data cannot be made freely available in a public repository.

/newpage

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.0     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter()     masks stats::filter()
## x dplyr::group_rows() masks kableExtra::group_rows()
## x dplyr::lag()        masks stats::lag()

##
## Attaching package: 'huxtable'

## The following object is masked from 'package:dplyr':
##
##     add_rownames
```