English Summary

Given the large burden of dementia affecting the population worldwide, major research efforts are taking place to identify ways to prevent and delay the onset of dementia. Several potential targets of intervention have been proposed, mostly based on large observational data, though findings have been sometimes contradictory. The aim of this thesis was to study the effect of different potential targets of intervention related to dementia prevention that have had controversial results in previous studies. To this matter, I overcame previous methodological challenges and frequent sources of bias by implementing causal inference theory and corresponding methods, such as target trial emulation and application of methods that take into account time-varying confounding feedback. To do so, I used data collected for the Rotterdam Study, a population-based cohort with rich longitudinal data assessment for over 20 years.

In **Chapter 2** I emulated a hypothetical randomized trial for estimating the intention-to-treat ("initiating treatment" vs. "not initiating treatment") and per-protocol effect ("initiating and sustained use" vs "not initiating ever") of statins in the risk of dementia. Eligibility criteria were defined in a way that allowed us to set a clear time zero, I conceptualized a "sequence of trials", which allowed us to assess eligibility criteria every month between 1993 and 2007, each of them with a 1-month enrollment period, increasing sample size and precision of estimates. In this study, I found that individuals with sustained statin use, but not statin initiation alone, had reduced 10-year risks of dementia and dementia or death. These results highlight the need to ask and answer questions related to a per-protocol effect.

Likewise, in **Chapter 3** I emulated a target trial to estimate the sustained effect of hypothetical interventions on systolic blood pressure (SBP) control, including in combination with an intervention on smoking cessation over follow-up, on the risk of first-ever stroke and dementia over 15 years of follow-up. By leveraging longitudinal information, and with the application of the parametric g-formula, I was able to simulate multiple treatment arms from the same source of data. All interventions were associated with a stroke risk reduction. In contrast, I did not observe a change in the risk of dementia, and all point estimates were above one. These results need to be interpreted in the context of death as a competing event. Given that I targeted a total effect, part of the effect on the risk of dementia is mediated by how interventions reduce the risk of death.

Since "reducing blood pressure" does not represent a well-defined intervention, in **Chapter 3** I discuss the different interpretations of the consistency assumption and limitations of our results. However, this study shows how, even when we don't have measures of the specific intervention, we can still conceptualize the target trial, under clear and transparent assumptions. In **Chapter 4**, I take this reasoning further to disentangle the potential sources of bias that could explain the inverse association between cancer and dementia found in previous studies. To this matter, I bring the Pin-1 hypothesis as part of the research question and progressively build a causal directed acyclic graph, outlining the assumptions needed to study the effect. I highlight the challenges that arise, which may introduce bias, and describe how these can be prevented (up to a certain extent) through different analytic decisions. Results show how depending on the confounding and selection bias control, and how cancer was used as a time-varying proxy, the risk ratio ranges from below the null to above the null.

Death played the role of a competing event in the previous studies. Since assumptions, results, and interpretation change depending on how we define this event, I aimed to understand the current practices on this topic in dementia research studies. In **Chapter 5** I present a systematic review of longitudinal studies focused (implicitly or explicitly) on causal effects in dementia risk. I found that almost half of the studies did not describe how death was handled in the methods section and only about ten percent had a clear and complete description of how death was treated in the main analysis. The vast majority presented estimates of a hazard ratio, mostly under a Cox proportional hazards model, though none reported the correct interpretation given the presence of a competing event nor discussed the related assumptions.

These results highlight the need for more educational applied resources in the area of competing events in causal inference. Thus, in **Chapter 6** I go through the key concepts of traditional estimands for competing events, the total effect and controlled direct effect, and outline their assumptions and interpretations. I present a hypothetical randomized trial on smoking cessation in late-midlife and emulate such a trial. The results highlight how different causal contrasts can result in different estimates, here going in opposite directions. Thus, we cannot begin to describe "bias" due to a competing event, let alone do something about that supposed bias, without stating clearly what question we were seeking an answer for.

In **Chapter 7** I discuss the findings and methodological implications of this dissertation while reflecting on the broader implications of the research and future areas of research. To conclude, in **Chapter 8** I reflect upon how the SARS-CoV-2 pandemic has forged my motivations and the next steps ahead.