**Towards a clearer causal question underlying the association between cancer and dementia**

## Supplemental digital content

## 

### Modeling description

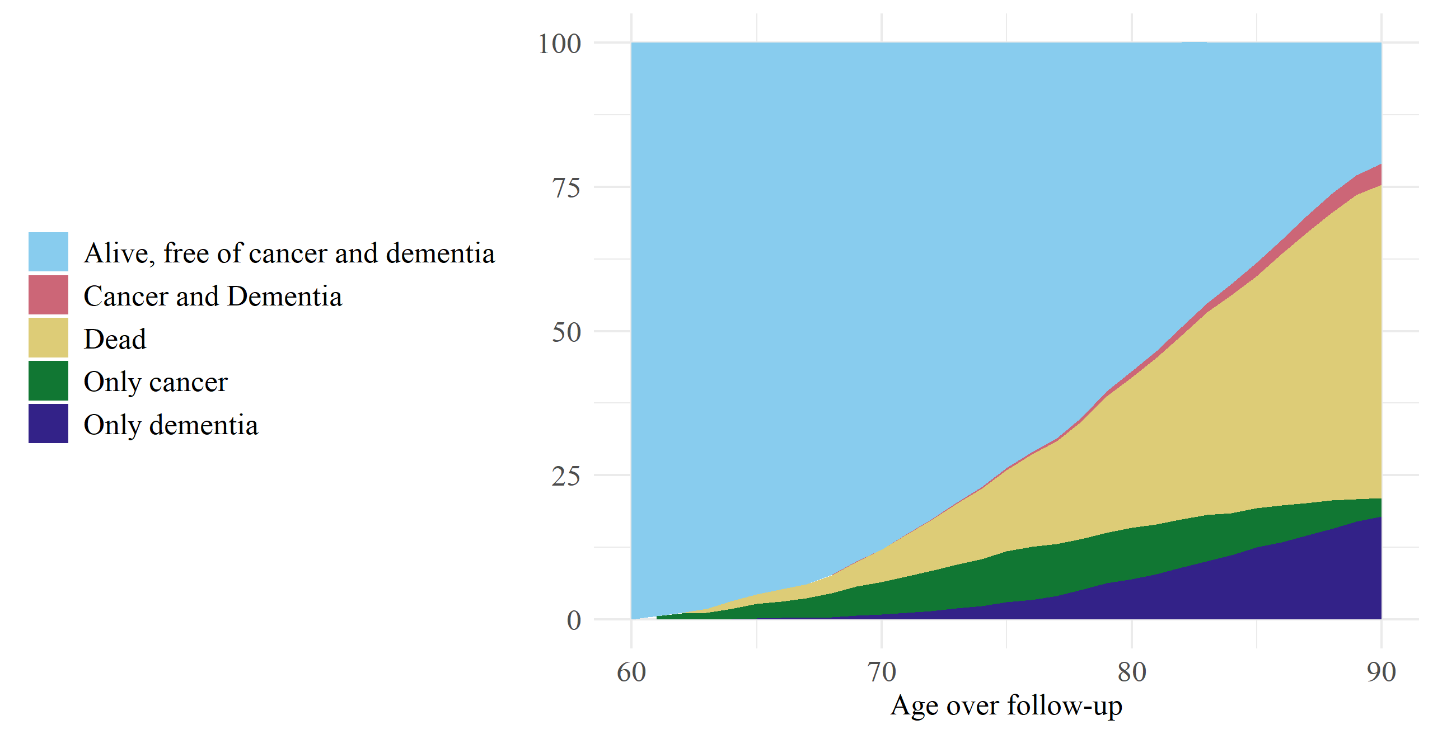
**Scenario A: Cancer diagnosis as “ever vs. never” proxy for Pin1**

|  |  |  |
| --- | --- | --- |
| **Weights** | **Numerator** | **Denominator** |
| **Inverse probability of treatment weights** | Pr[At+k = 1|1]  Where:  *A* = cancer diagnosis (0 = never, 1 = ever) | Pr[At+k = 1|V]  Where:  *A*= cancer diagnosis (0 = never, 1 = ever); *V* = age at study entry with natural cubic splines, sex (women vs. men), education (three categories), apoe4 (three categories), smoking status at study entry(three categories), cohort (two categories); product terms between covariates |
| **Inverse probability of censoring weights for death** | Pr[Dt+k = 0| Yt+k = 0, 1]  Where:  *D* = died prior to dementia diagnosis over follow-up (0 = no, 1 = yes); Y = dementia diagnosis | Pr[Dt+k = 0|Yt+k = 0, A t+k,L]  Where:  *D* = died prior to dementia diagnosis over follow-up (0 = no, 1 = yes); Y = dementia diagnosis (0 = no, 1 = yes); *A* = cancer diagnosis (0 = never, 1 = ever); *V* = age at study entry with natural cubic splines, sex (women vs. men), education (three categories), apoe4 (three categories), cohort (two categories), smoking status at study entry(three categories), hypertension status at study entry (two categories), history of diabetes at study entry (three categories), systolic blood pressure and body mass index at study entry (continuous variables); product terms between covariates |

**Scenario B. Cancer diagnosis as a time-varying proxy for Pin1**

|  |  |  |
| --- | --- | --- |
| **Inverse probability of treatment weights** | Pr[At+1 = 1| At = 0, Dk = 0, Yk = 0, T]  Where:  *A*= cancer diagnosis (0 = no, 1 = yes); *T* = year with natural cubic splines | Pr[A*t+1* = 1| At = 0, Dt = 0, Yt = 0, T, V, Lt]  Where *A* = cancer diagnosis (0 = no, 1 = yes); *T* = year with natural cubic splines; *V* = age at study entry with cubic splines, sex (women vs. men), education (three categories), apoe (three categories), cohort (two categories); *L*= time-varying smoking status (three categories), hypertension status (two categories), hypertension medication (two categories), diabetes status (three categories), systolic blood pressure and body mass (continuous variables); no product terms between covariates |
| **Inverse probability of censoring weights for death** | Pr[D*t+1* = 0| D*t* = 0, Y*t* = 0, A*t*, T, V]  Where:  *D* = death (0 = no, 1 = yes); *Y =* dementia diagnosis (0 = no, 1 = yes); *A* = cancer diagnosis (0 = no, 1 = yes), *Y* = year with natural cubic splines; *V* = cohort (two categories); no product terms between covariates | Pr[D*k+1* = 0| D*k* = 0, Y*t* = 0, A*t*, T, V, L*t*]  Where *D* = death (0 = no, 1 = yes); *Y* = dementia diagnosis (0 = no, 1 = yes); ; *A*= cancer diagnosis (0 = no, 1 = yes); *T =* year with natural cubic splines; *V* = age at study entry with cubic splines, sex (women vs. men), education (five categories), apoe (three categories), cohort (two categories); *L*t = time-varying smoking status (three categories), hypertension status (two categories), hypertension medication (two categories), diabetes status (three categories), heart disease condition (yes, no), incident stroke (yes, no), systolic blood pressure and body mass (continuous variables) and no product terms between covariates |

### eFigure 1. Distribution of participants under each health status, by age over follow-up



### eFigure 2. Causes of death defined by ICD10 codes for participants who died prior to dementia diagnosis

C:\Users\040609\AppData\Local\Microsoft\Windows\INetCache\Content.Word\causes_death.tiff

Information on vital status and cause – specific mortality is obtained continuously from the municipal authorities in Rotterdam. General practitioners in the research area report incident events by means of a computerized system, covering 78.8 % of the cohort. General practitioners without the computerized system are requested to notify new events annually. Trained research assistants subsequently collect information from medical records of the general practitioners, hospitals, and nursing homes. Two research physicians independently classify the events according to the ICPC and ICD-10 coding systems. Thereafter, a medical expert in the field reviews all coded events.