Association between cancer and dementia risk in the UK Biobank: evidence of diagnostic bias

Jingxuan Wang[1], Peter Buto[1], Sarah F. Ackley[1], Lindsay C. Kobayashi[2], Rebecca E. Graff[1], Scott C. Zimmerman[1], Stephen B. Asiimwe[1], Camilla Calmasini[1], M. Maria Glymour[1]

1. Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA
2. Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA

Key words: dementia, Alzheimer's disease, cancer, diagnostic bias

Abstract

Introduction:

Epidemiological studies have identified an inverse association between cancer and dementia. Underlying methodological biases such as confounding, diagnostic bias, and selective survival have been postulated, yet no studies have systematically investigated the potential for each source of bias within a single dataset. We used the UK Biobank to generate and compare estimates for the cancer-dementia association using different analytical specifications designed to address different sources of bias.

Methods:

We included 140,959 UK Biobank participants aged ≥55 without dementia before enrollment, and with linked primary care data. We used cancer registry data to identify cases of prevalent cancer before UK Biobank enrollment and incident cancer after enrollment. We used multivariable-adjusted Cox models to evaluate the associations of prevalent and incident cancer with incidence of all-cause dementia, Alzheimer's disease (AD), and vascular dementia. And we systematically evaluated each potential source of bias.

Results:

The cohort accumulated 3,310 incident dementia diagnoses over a median of 12.3 years of follow-up. All-site incident cancer was positively associated with all-cause dementia risk (hazard ratio [HR]=1.14, 95% CI: 1.02-1.29). The adjusted HR for prevalent cancer was 1.03 (95% CI: 0.92-1.17). Results were similar for vascular dementia. AD dementia was not associated with prevalent or incident cancer. Dementia diagnosis was substantially elevated in the first year after cancer diagnosis (HR=1.83, 95% CI: 1.48-2.36), suggesting diagnostic bias.

Conclusion:

Incident cancer diagnoses were associated with a higher risk of subsequent dementia diagnoses. Increased chance of dementia diagnosis associated with increased health care utilization after a cancer diagnosis may be a source of bias in electronic health records-based studies.

Introduction

Cancer and dementia are leading causes of death globally, and aging is related to increased risks of both diseases [1–3]. Several epidemiological studies have reported an inverse association between cancer and dementia, particularly AD [4–8]. In a recent meta-analysis of 22 cohort and case-control studies, an incident cancer diagnosis was associated with an 11% reduction in the subsequent risk of AD [9]. A number of biological mechanisms have been proposed to explain why individuals with a history of cancer may have a lower risk of dementia, including differential cell regulation of senescence and proliferation[10,11], shared genetic pathways through the *PIN1* and *PARK2* genes [12,13], and chronic inflammation and immunosenescence [14].

However, epidemiological results are inconsistent, and several methodological biases have been postulated to explain the findings [9]. In a recent review, Ospina-Romero et al. identified the potential for bias due to confounding, diagnosis, selective survival, and competing risk of death. Confounding may account for the inverse relationship in analyses that did not adequately adjust for extraneous common causes of cancer and dementia. Diagnostic bias may emerge when cancer diagnoses and treatments cause clinicians to overlook dementia symptoms, delaying the diagnosis of dementia among cancer patients. However, diagnostic bias may also work in the opposite direction as cancer patients have more frequent contact with clinicians, increasing their chance of dementia diagnosis [15]. Survival bias may arise if there are unmeasured protective characteristics of individuals with cancer, which can promote their survival while reducing their dementia risk relative to cancer-free individuals [16]. Finally, the competing risk of death may introduce bias because people with cancer have a higher mortality risk than cancer-free individuals, and may not live long enough for the diagnosis of dementia [13,15].

To better understand what factors may drive the association between cancer and dementia, we used longitudinal data from the UK Biobank to investigate the risks of all-cause dementia, AD, and vascular dementia among survivors of prevalent and incident cancers. We leveraged features of the data to evaluate several potential spurious explanations for a cancer-dementia association. Specifically, we assessed 1) confounding through specification of three sets of covariates, 2) diagnostic bias by adjusting for the frequency of primary care contact as a proxy for medical surveillance, 3) survival bias by examining the association between non-melanoma skin cancer (NMSC) and dementia because this type of cancer does not substantially increase mortality risk [17], and 4) competing risk by using transport injuries as a negative control outcome [18].

Methods

Study population

The UK Biobank is a large, prospective cohort study of 502,411 adults aged 40-69 years were recruited from 2006-2010[19]. Participants attended one of 22 assessment centers across England, Wales, and Scotland, where they completed a touchscreen questionnaire and a face-to-face interview with study nurses. Participants also underwent physical measurements and provided blood, urine, and saliva samples. Health outcomes were measured in lined National Health Service (NHS) hospital admissions data, national death and cancer registers, and primary care data for more than 230,000 participants. For these analyses, we included participants who were 55 years or older and free of dementia at the baseline UK Biobank assessment, and who had primary care data linked to the UK Biobank database (Figure 1).

Identification of cancer cases
We identified cancer cases and date of diagnosis using data from cancer registries (NHS Digital and Public Health England for England and Wales; NHS Central Register for Scotland) and self-report (sensitivity analyses only, described below). In our primary analysis, we included all cancer types except NMSC. We also considered the following site-specific cancers: NMSC, breast cancer, and prostate cancer. We treated cancer diagnosis as a time-varying exposure and categorized participants as having 1) prevalent cancer if there was a cancer diagnosis in their medical record prior to the date of UK Biobank enrollment, 2) incident cancer if there was a cancer diagnosis in their medical record after their date of UK Biobank enrollment, or 3) no cancer if there was no record of cancer in their medical records, either prior to or after enrollment. All-cause cancer diagnoses were defined as an ICD-10 code of C00-C96 except C44 (NMSC), D00-D49, or an ICD-9 code of 140-239 except 173 (NMSC).

Ascertainment of dementia cases
The primary outcome for our analysis was incident all-cause dementia diagnosis. All-cause dementia was ascertained via linked NHS hospital admissions, mortality, and primary care data. In the UK Biobank, hospital inpatient data and death records are coded using ICD-9 and ICD-10, and primary care records are coded using Read version 2 (Read v2) and Clinical Terms Version 3 (CTV3), both of which are part of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT).We used a comprehensive list of ICD codes to identify all-cause dementia cases, including AD, vascular dementia, frontotemporal dementia, Lewy body dementia, alcohol-related dementia, and Creutzfeldt-Jakob disease (Supplementary Table 6). Incident dementia was defined as having dementia as a primary or secondary diagnosis in hospital inpatient or primary care data, or a contributory cause of death in mortality data. In addition to all-cause dementia, we evaluated the incidence of AD and vascular dementia as separate outcomes.

Covariates
The following covariates were included in multivariable analyses: age, sex, race, assessment center, the presence of the APOE-ε4 allele, education, Townsend deprivation index, number of clinic visits in the year before study entry, body mass index (BMI), ever smoked, ever used alcohol, physical activity, diabetes, and stroke. Age and sex were acquired from a central registry at recruitment and updated by participants. Race, education, ever smoked, and ever used alcohol were self-reported by participants through a touchscreen questionnaire at baseline. Townsend deprivation index was calculated prior to baseline based on preceding national census output areas and determined by participants' postcodes at recruitment. The number of clinic visits in the year before study entry was determined by the number of days with clinical interactions in the primary care data. BMI was calculated from height and weight that were measured by staff during the baseline visit and treated as a continuous variable. Physical activity was defined as self-report of ≥150 min/week of moderate activity or ≥75 min/week of vigorous activity or an equivalent combination [20]. Diabetes and stroke histories at baseline were identified via linked inpatient, primary care, mortality data, or self-reported medical conditions. For analyses, we considered three sets of covariates. Model 1 included sex, race, assessment center, and the presence of the APOE-ε4 allele. Model 2 additionally included education, Townsend deprivation index, and the number of clinic visits. Model 3 additionally included BMI, ever smoked, ever used alcohol, physical activity, diabetes, and stroke.

Statistical analyses

We summarized the baseline characteristics of the study sample stratified by cancer status (no diagnosis, prevalent at baseline, or incident during follow-up). To examine the relationship between a cancer diagnosis and an incident dementia diagnosis, we used Cox proportional hazards regression models. Participants were followed from baseline to the date of first dementia diagnosis, death from any cause, or the latest date at which hospital inpatient data were available (September 30, 2021 for England; July 31, 2021 for Scotland; and February 28, 2018 for Wales), whichever came first. We used age as the time scale because mortality and morbidity risks are expected to increase with age. We conducted all analyses stratified by baseline age categories (55-59, 60-64, 65-69, ≥70 years), to allow the underlying baseline hazard function to vary by age category. In Models 2 and 3, we modelled the number of clinic visits flexibly with restricted cubic splines. Cox proportional hazards regression models were stratified by assessment center (all models) and ever smoked (model 3).

In our primary analysis, participants with an incident cancer diagnosis contributed person-time to the "no cancer" group until their date of cancer diagnosis, after which they contributed to person-time in the incident cancer group. In a secondary analysis, to account for potential selective survival and ascertainment bias by cancer status, we calculated the hazard ratios for dementia over five time intervals after the earliest cancer diagnosis (0 – 1 year, >1 – 5 years, >5 – 10 years, >10 – 20 years, and >20 years after cancer diagnosis) compared to individuals with no prior cancer diagnosis.

In order to assess competing risk, we repeated the above analysis using transportation injuries as a negative control. We chose transportation injuries because there is no known biological mechanism in the relationship between transportation injuries and cancer. In the presence of possible bias from competing risk due to death, mortality due to cancer precludes dementia onset and we expect to observe a negative association between cancer and transportation injuries.

We assessed violations of the proportional hazards assumption with Schoenfeld residual tests. We repeated the analyses using the secondary outcomes (AD and vascular dementia). In addition, we analyzed site-specific associations between cancer and all-cause dementia for breast, prostate, and NMSC, as cancers at each of these sites have been linked to dementia in previous research[17,21–23].

We conducted three sensitivity analyses to evaluate robustness of our results. First, we examined the impact of potential incompleteness in the health record data on cancer diagnoses by including data on self-reported cancer history[24]. Our second sensitivity analysis assessed the potential impact of missing data using multiple imputation by chained equations with 10 imputations[25]. The percentage of missing variables ranged from 0% for sex, age, assessment center, number of clinic visits, and comorbidities, to 2.7% for self-reported physical activity. Most covariates had less than 1% of values missing. Finally, we repeated the primary analyses restricting to individuals with no prevalent cancer diagnosis at study enrollment and evaluated the association of incident cancer with dementia diagnosis. All statistical analyses were performed using R version 4.0.5.

Results

Among 140,959 participants who met the inclusion criteria, 12,635 (9.0%) had prevalent cancer diagnosed before baseline, and 17,905 (12.7%) had an incident cancer diagnosed during follow-up. Over a median follow-up of 12.3 years (interquartile range, 11.3-13.1 years), corresponding to 1,671,906 person-years at risk for dementia, there were 3,310 cases of incident all-cause dementia. Baseline characteristics stratified by cancer status are shown in Table 1. The analytical sample was predominantly White (96.8%). The median baseline age of both prevalent and incident cancer patients was 63.4 years, which was slightly older than cancer-free individuals (62.2 years). Dementia incidence rates increased almost exponentially with age, regardless of cancer status (Figure 2).

In the fully adjusted model (Table 2, Model 3), prevalent cancer was not significantly associated with all-cause dementia (HR = 1.03, 95% CI: 0.92-1.17), while incident cancer was significantly associated with a higher risk of all-cause dementia (HR = 1.14, 95% CI: 1.02-1.29). A similar hazard ratio is observed for vascular dementia for both prevalent cancer (HR = 1.13, 95% CI: 0.90-1.17) and incident cancer (HR = 1.27, 95% CI: 1.02-1.58; Table 2, Model 3), although not statistically significant for prevalent cancer. Prevalent and incident cancers were not significantly associated with AD. In the sensitivity analysis including self-reported cancer diagnoses (Supplementary Table 3), results were similar (HR for all-cause dementia = 1.03, 95% CI: 0.92-1.17 for prevalent cancer; HR = 1.14, 95% CI = 1.02-1.29 for incident cancer). Using multiple imputation to impute missing covariate values (Supplementary Table 4) or excluding individuals with prevalent cancer diagnosis at study enrollment (Supplementary Table 5) did not substantially change our results. Diagnoses of dementia were highest in the first year after cancer diagnosis compared to individuals with no cancer diagnosis (HR = 1.83, 95% CI = 1.48-2.36), then the association attenuated (HR = 1.03, 95% CI = 0.87-1.22 during 1-5 years after cancer diagnosis; HR = 1.04, 95% CI = 0.88-1.23 during 5-10 years after cancer diagnosis). Similar patterns were observed for AD and vascular dementia (see Table 3).

No significant associations between cancer and risk of subsequent transportation injuries were found (Supplementary Table 1: HR = 0.96, 95% CI: 0.78-1.18 for prevalent cancer; HR = 0.93, 95% CI: 0.74-1.19 for incident cancer).

Diverging associations were observed for site-specific cancer diagnoses (Supplementary Table 2). In the fully adjusted model, incident NMSC diagnosis was associated with a significantly lower risk of all-cause dementia (HR = 0.82, 95% CI: 0.67-1.00). There were no significant associations between either prevalent or incident prostate cancer diagnoses and all-cause dementia among males. Similarly, among females, no significant associations between breast cancer and all-cause dementia were found.

Discussion

In this large-scale prospective cohort study, we found a significant positive association between overall cancer diagnoses and subsequent all-cause dementia and vascular dementia. These results were driven primarily by a large elevation in the probability of a dementia diagnosis in the year immediately following a new cancer diagnosis. No significant association between overall

prevalent or incident cancer diagnoses and AD was found. Incident nonmelanoma skin cancer diagnosis was associated with a lower risk of all-cause dementia, but prostate and breast cancer were not.

Our results are generally consistent with those from previous studies that ascertained dementia diagnoses from electronic health records [9]. Our finding of a temporal trend in the cancer-dementia association suggests that the increased probability of dementia diagnosis in the first year after a cancer diagnosis may be an important source of bias in electronic health records-based studies of cancer and dementia. In Ospina-Romero et al.'s meta-analysis, the pooled HR for studies that ascertained AD from electronic health records was 0.94 (95% CI: 0.58-1.52). A recent population-based cohort study of US veterans found an increased risk of subsequent non-AD dementia among survivors of all cancers (HR = 1.17; 95% CI: 1.15-1.19) and no association between all cancers and AD (HR = 1.00; 95% CI: 0.97-1.03) [26]. On the other hand, in Ospina-Romero et al.'s meta-analysis, community-based studies that did not use medical records to ascertain dementia status and thus classified as less susceptible to AD diagnostic bias showed an inverse association between cancer and AD (HR = 0.73; 95% CI: 0.58-0.90) [9]. For example, a community-based cohort study reported mixed evidence for the association between all cancers and subsequent dementia among survivors of prevalent cancer before study entry (HR = 0.92; 95% CI: 0.76-1.11) and incident cancer during follow-up (HR = 0.87; 95% CI: 0.64-1.04)[27].

Prior work suggests a possible inverse biological link between the risk of neoplastic disease and risk of neurodegenerative diagnosis. Our findings do not support this possibility, rather suggesting that diagnostic bias may lead to the observation of inverse associations in EHR-based studies. The latter possibility is supported in our analysis by a spike of dementia diagnoses immediately following cancer diagnosis.

Strengths and limitations

The primary limitations of this study include the limited number of dementia cases, potential for diagnostic bias due to the use of electronic health records, and over-representation of healthy and predominantly White individuals with high socioeconomic status (SES) in the UK Biobank. Because this cohort is relatively young, with a mean age of 74.4 years (SD=4.3) at the end of follow-up, later-onset dementia would not have been observed. Dementia incidence rates in our sample were substantially lower than in the general UK population [28]. The limited number of incident dementia cases may also have been due to a healthy volunteer selection bias in the UK Biobank [29]. Unlike community-based cohort studies in which routine dementia screening for all participants is completed at predetermined intervals, our study may be vulnerable to diagnostic bias because we derived dementia cases from electronic health records. We attempted to minimize the diagnostic bias by adjusting for the number of primary care clinics visits in the year before study entry. Comparison between diagnostic codes of all-cause dementia and clinical expert adjudication for a subset of UK Biobank participants showed a positive predictive value of 82.5%[30], suggesting some misclassification of dementia diagnosis.

The major strengths of this cohort study include the prospective design, large sample size, long follow-up period, and its rich data that allow the assessment of potential methodological biases. Detailed individual information on demographics, behaviors, and comorbidities were collected at baseline, which enabled us to control for multiple potential confounders. We also adjusted for the

frequency of primary care clinic visits in order to account for potentially differential health care utilization by cancer status. Survival bias due to differential mortality in the cancer and cancer-free groups may have been avoided by separating prevalent and incident cancer, and separately investigating NMSC, which does not substantially increase mortality [17]. In addition, our analysis of transportation injuries after a cancer diagnosis did not suggest differential mortality between cancer and cancer-free groups.

Conclusion and future work

We observed a significantly higher risk of dementia among incident cancer patients than their cancer-free counterparts among adults aged ≥55 in the UK Biobank. Many of these dementia cases were diagnosed in the first year following a cancer diagnosis, suggesting that increased health care utilization following a cancer diagnosis may increase the likelihood of dementia ascertainment. Hence, diagnostic bias could contribute to or explain the association between cancer and dementia observed in electronic health records-based studies. Future studies of the relationship between cancer and dementia should address these methodological biases to improve causal inference and expand evidence on the etiologies of cancer and dementia. Studies that leverage different methodological approaches, such as evaluating genetic linkage and using neuroimaging biomarkers, may be required to fully understand the relationship between cancer and dementia.

References:

1. World Health Organization. Global health estimates: Leading causes of death [Internet]. Available from: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death

2. Scheltens P, Strooper BD, Kivipelto M, Holstege H, Chételat G, Teunissen CE, et al. Alzheimer's disease. The Lancet. Elsevier; 2021;397:1577–90.

3. Soerjomataram I, Bray F. Planning for tomorrow: global cancer incidence and the role of prevention 2020–2070. Nat Rev Clin Oncol. Nature Publishing Group; 2021;18:663–72.

4. Driver JA, Beiser A, Au R, Kreger BE, Splansky GL, Kurth T, et al. Inverse association between cancer and Alzheimer's disease: results from the Framingham Heart Study. BMJ. British Medical Journal Publishing Group; 2012;344:e1442.

5. Roe CM, Fitzpatrick AL, Xiong C, Sieh W, Kuller L, Miller JP, et al. Cancer linked to Alzheimer disease but not vascular dementia. Neurology. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology; 2010;74:106–12.

6. Musicco M, Adorni F, Santo SD, Prinelli F, Pettenati C, Caltagirone C, et al. Inverse occurrence of cancer and Alzheimer disease: A population-based incidence study. Neurology. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology; 2013;81:322–8.

7. Freedman DM, Wu J, Chen H, Kuncl RW, Enewold LR, Engels EA, et al. Associations between cancer and Alzheimer's disease in a U.S. Medicare population. Cancer Med. 2016;5:2965–76.

8. Sun M, Wang Y, Sundquist J, Sundquist K, Ji J. The Association Between Cancer and Dementia: A National Cohort Study in Sweden. Front Oncol [Internet]. Frontiers; 2020 [cited 2021 May 27];10. Available from: https://www.frontiersin.org/articles/10.3389/fonc.2020.00073/full

9. Ospina-Romero M, Glymour MM, Hayes-Larson E, Mayeda ER, Graff RE, Brenowitz WD, et al. Association Between Alzheimer Disease and Cancer With Evaluation of Study Biases: A Systematic Review and Meta-analysis. JAMA Netw Open. 2020;3:e2025515.

10. Behrens MI, Lendon C, Roe CM. A common biological mechanism in cancer and Alzheimer's disease? Curr Alzheimer Res. 2009;6:196–204.

11. Holohan KN, Lahiri DK, Schneider BP, Foroud TM, Saykin AJ. Functional microRNAs in Alzheimer's disease and cancer: differential regulation of common mechanisms and pathways. Front Genet [Internet]. Frontiers; 2013 [cited 2021 May 28];3. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2012.00323/full

12. Li Q, Dong Z, Lin Y, Jia X, Li Q, Jiang H, et al. The rs2233678 Polymorphism in PIN1 Promoter Region Reduced Cancer Risk: A Meta-Analysis. PLOS ONE. Public Library of Science; 2013;8:e68148.

13. Driver JA. Inverse association between cancer and neurodegenerative disease: review of the epidemiologic and biological evidence. Biogerontology. 2014;15:547–57.

14. Ganguli M. Cancer and Dementia: It's Complicated. Alzheimer Dis Assoc Disord. 2015;29:177–82.

15. Willik KD van der, Schagen SB, Ikram MA. Cancer and dementia: Two sides of the same coin? Eur J Clin Invest. 2018;48:e13019.

16. Hayes-Larson E, Ackley SF, Zimmerman SC, Ospina-Romero M, Glymour MM, Graff RE, et al. The competing risk of death and selective survival cannot fully explain the inverse cancer-dementia association. Alzheimers Dement J Alzheimers Assoc. 2020;16:1696–703.

17. Schmidt SAJ, Ording AG, Horváth-Puhó E, Sørensen HT, Henderson VW. Non-melanoma skin cancer and risk of Alzheimer's disease and all-cause dementia. PLOS ONE. Public Library of Science; 2017;12:e0171527.

18. Arnold BF, Ercumen A, Benjamin-Chung J, Colford JM. Brief Report: Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies. Epidemiol Camb Mass. 2016;27:637–41.

19. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. Health Policy Technol. 2012;1:123–6.

20. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, et al. Defining and Setting National Goals for Cardiovascular Health Promotion and Disease Reduction. Circulation. American Heart Association; 2010;121:586–613.

21. Ording AG, Horváth-Puhó E, Veres K, Glymour MM, Rørth M, Sørensen HT, et al. Cancer and risk of Alzheimer's disease: Small association in a nationwide cohort study. Alzheimers Dement. 2020;16:953–64.

22. Robinson D, Garmo H, Van Hemelrijck M, Damber J, Bratt O, Holmberg L, et al. Androgen deprivation therapy for prostate cancer and risk of dementia. Bju Int. 2019;124:87–92.

23. Sun L-M, Chen H-J, Liang J-A, Kao C-H. Long-term use of tamoxifen reduces the risk of dementia: a nationwide population-based cohort study. QJM Int J Med. 2016;109:103–9.

24. Møller H, Richards S, Hanchett N, Riaz SP, Lüchtenborg M, Holmberg L, et al. Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates. Br J Cancer. 2011;105:170–6.

25. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30:377–99.

26. Frain L, Swanson D, Cho K, Gagnon D, Lu KP, Betensky RA, et al. Association of cancer and Alzheimer's disease risk in a national cohort of veterans. Alzheimers Dement J Alzheimers Assoc. 2017;13:1364–70.

27. Bowles EJA, Walker RL, Anderson ML, Dublin S, Crane PK, Larson EB. Risk of Alzheimer's disease or dementia following a cancer diagnosis. PloS One. 2017;12:e0179857.

28. Pham TM, Petersen I, Walters K, Raine R, Manthorpe J, Mukadam N, et al. Trends in dementia diagnosis rates in UK ethnic groups: analysis of UK primary care data. Clin Epidemiol. 2018;10:949–60.

29. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017;186:1026–34.

30. Wilkinson T, Schnier C, Bush K, Rannikmäe K, Henshall DE, Lerpiniere C, et al. Identifying dementia outcomes in UK Biobank: a validation study of primary care, hospital admissions and mortality data. Eur J Epidemiol. 2019;34:557–65.

Figure 1: Flow diagram of inclusion criteria

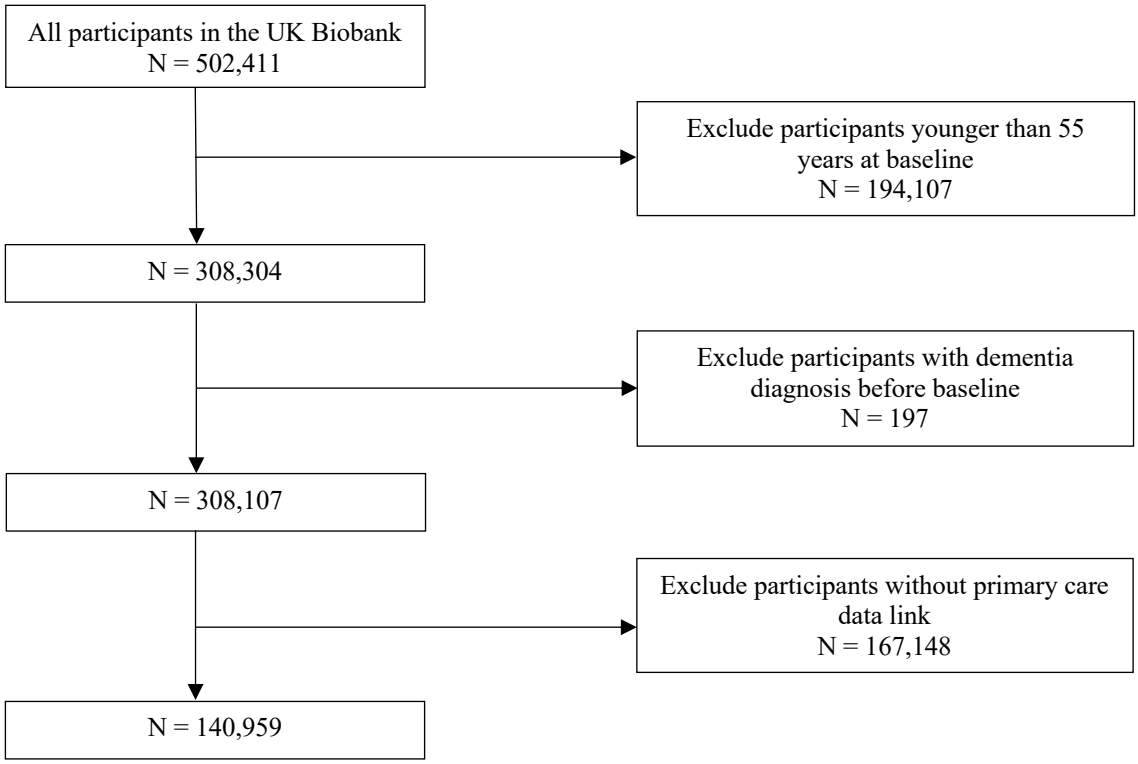
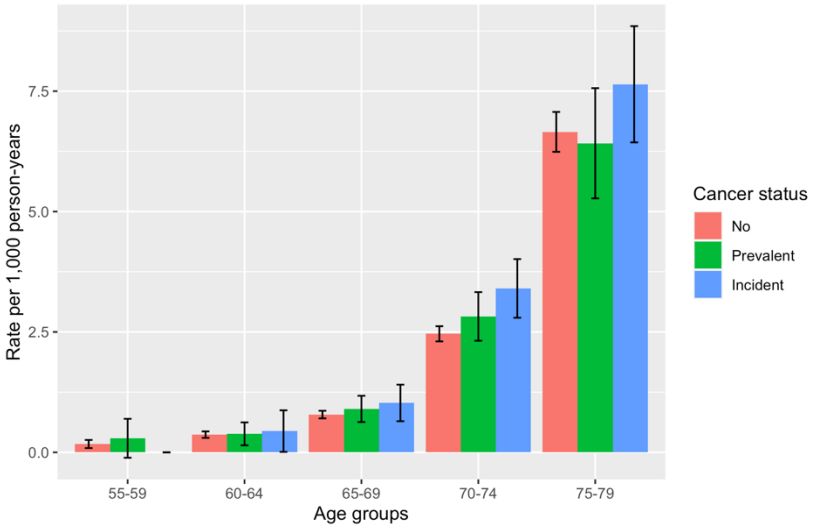


Figure 2: Age-specific all-cause dementia incidence rates



Age-specific incidence rates were calculated by dividing the number of all-cause dementia cases in each age group by the number of person-years of observation in that group. Participants were at risk and contributed person-years from study baseline until the time of first dementia diagnosis, death from any cause, or censoring date. For each age group, a 95% CI for the age-specific incidence rate was calculated assuming a Poisson distribution for the number of cases in that group.

Table 1: Baseline characteristics of analytical sample stratified by cancer status

Characteristic	No cancer No. (%)[a]	Prevalent cancer No. (%)[a]	Incident cancer No. (%)[a]
Number	110,419	12,635	17,905
Median age (IQR)	62.2 (59.0-65.6)	63.4 (60.2-66.6)	63.4 (60.4-66.5)
Sex			
Female	59,941(54.3)	8,299 (65.7)	7,859 (43.9)
Male	50,478(45.7)	4,336 (34.3)	10,046 (56.1)
Race			
White	106,516(96.6)	12,356 (97.8)	17,394 (96.9)
Black	718 (0.7)	69 (0.5)	117 (0.7)
Asian	2,021 (1.8)	101 (0.8)	219 (1.2)
Other	1,021 (0.9)	102 (0.8)	152 (0.9)
Education			
Higher	38,723 (35.1)	4,277 (33.9)	5,932 (33.2)
Secondary	35,883 (32.6)	4,079 (32.3)	5,570 (31.2)
Vocational	8,018 (7.3)	857 (6.8)	1,428 (8.0)
Other	27,610 (25.0)	3,413 (27.0)	4,944 (27.7)
BMI, mean (SD)	27.6 (4.7)	27.6 (4.8)	29.0 (4.7)
Physical activity			
Yes	84,143(78.3)	9,323 (76.1)	13,500 (77.5)
No	23,315 (21.7)	2,934 (23.9)	3,915 (22.5)
Ever smoked			
Yes	52,041 (47.4)	6,451 (51.4)	9,690 (54.5)
No	57,757 (52.6)	6,107 (48.6)	8,102 (45.5)
Ever used alcohol			
Yes	104,978 (95.3)	11,991 (95.0)	17,154 (96.1)
No	5,148 (4.7)	625 (5.0)	700 (3.9)
Townsend deprivation index[b]			
1 (least deprived)	23,221 (21.1)	2,576 (20.4)	3,681 (20.6)
2-4	68,199 (61.8)	7,756 (61.4)	10,984 (61.4)
5 (most deprived)	18,859 (17.1)	2,291 (18.2)	3,226 (18.0)
Diabetes			
Yes	6,952 (6.3)	843 (6.7)	1,385 (7.7)
No	103,467 (93.7)	11,792 (93.3)	16,520 (92.3)
Stroke			
Yes	2,191 (2.0)	301 (2.4)	422 (2.4)
No	108,228 (98.0)	12,334 (97.6)	17,483 (97.6)
Median number of clinic visits[c] (IQR)	6 (2-11)	8 (3-14)	6 (2-12)
Median follow-up time (IQR)	12.3 (11.6-13.1)	12.2 (11.3-13.0)	4.9 (2.2-8.2)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); IQR, interquartile range; SD, standard deviation.

[a] Percentages may not sum to 100 due to rounding.
[b] Index quintiles, combining social class, employment, car availability, and housing.
[c] Number of primary care clinics visits in the year before study entry.

Table 2: Risk of incident all-cause dementia, AD, and vascular dementia after a cancer diagnosis

Outcome	No. of events /person-years[a]	Age-adjusted IR[b]	HR	95% CI	HR	95% CI	HR	95% CI
			Model 1[c]		Model 2[d]		Model 3[e]	
All-cause dementia								
No cancer	2,640/1,432,923	3.28	1.00	ref.	1.00	ref.	1.00	ref.
Prevalent cancer	327/143,335	3.54	1.09	0.97-1.22	1.02	0.90-1.15	1.03	0.92-1.17
Incident cancer	343/95,648	3.75	1.16	1.03-1.30	1.15	1.03-1.29	1.14	1.02-1.29
AD dementia								
No cancer	1,222/1,437,116	1.76	1.00	ref.	1.00	ref.	1.00	ref.
Prevalent cancer	139/143,834	1.50	0.96	0.80-1.14	0.92	0.77-1.11	0.95	0.79-1.14
Incident cancer	147/96,121	1.70	1.07	0.90-1.27	1.07	0.90-1.27	1.04	0.87-1.25
Vascular dementia								
No cancer	664/1,438,617	0.87	1.00	ref.	1.00	ref.	1.00	ref.
Prevalent cancer	91/144,010	0.73	1.18	0.94-1.48	1.08	0.86-1.36	1.13	0.90-1.43
Incident cancer	95/96,283	1.16	1.26	1.01-1.57	1.25	1.00-1.56	1.27	1.02-1.58

Abbreviations: AD, Alzheimer's disease; CI, confidence interval; HR, hazard ratio; IR, incidence rate.
[a] Number of dementia cases divided by follow-up time (person-years) in cancer group.
[b] Per 1,000 person-years.
[c] Adjusted for sex, race, and the presence of the APOE-ε4 allele, stratified by assessment center.
[d] Further adjusted for education, Townsend deprivation index, and the number of clinic visits.
[e] Further adjusted for BMI, ever used alcohol, physical activity, diabetes, and stroke, further stratified by ever smoked.

Table 3: Risk of incident all-cause dementia, AD, and vascular dementia during five time periods after cancer diagnosis

Outcome	No. of events /person-years[a]	HR	95% CI	HR	95% CI	HR	95% CI
		Model 1[b]		Model 2[c]		Model 3[d]	
All-cause dementia							
No cancer	2,640/1,432,923	1.00	ref.	1.00	ref.	1.00	ref.
0-1 years	65/16,960	1.83	1.42-2.35	1.79	1.39-2.30	1.83	1.48-2.36
1-5 years	147/60,253	1.04	0.88-1.24	1.03	0.87-1.22	1.03	0.87-1.22
5-10 years	163/57,353	1.10	0.93-1.29	1.06	0.90-1.24	1.04	0.88-1.23
10-20 years	190/63,426	1.10	0.95-1.28	1.04	0.89-1.21	1.04	0.89-1.21
>20 years	105/40,991	1.06	0.87-1.29	1.02	0.84-1.25	1.06	0.87-1.31
AD dementia							
No cancer	1,222/1,437,116	1.00	ref.	1.00	ref.	1.00	ref.
0-1 years	35/17,018	2.22	1.57-3.12	2.18	1.55-3.07	2.16	1.52-3.08
1-5 years	57/60,454	0.88	0.68-1.16	0.88	0.67-1.15	0.87	0.66-1.15
5-10 years	64/57,600	0.94	0.73-1.21	0.92	0.71-1.19	0.91	0.70-1.18
10-20 years	78/63,710	0.91	0.72-1.16	0.88	0.70-1.12	0.88	0.69-1.13
>20 years	52/41,172	1.06	0.80-1.41	1.04	0.78-1.38	1.10	0.82-1.46
Vascular dementia							
No cancer	664 /1,438,617	1.00	ref.	1.00	ref.	1.00	ref.
0-1 years	17/17,036	1.98	1.22-3.21	1.94	1.19-3.14	2.08	1.28-3.38
1-5 years	38/60,554	1.10	0.79-1.52	1.07	0.77-1.49	1.13	0.81-1.57
5-10 years	55/57,664	1.41	1.07-1.87	1.35	1.02-1.78	1.30	0.97-1.75
10-20 years	43/63,830	0.98	0.72-1.34	0.91	0.66-1.24	0.96	0.70-1.32
>20 years	33/41,209	1.26	0.87-1.83	1.20	0.83-1.75	1.27	0.87-1.87

Abbreviations: AD, Alzheimer's disease; CI, confidence interval; HR, hazard ratio.

[a] Number of dementia cases divided by follow-up time (person-years) in the cancer group.

[b] Adjusted for sex, race, and the presence of the APOE-ε4 allele, stratified by assessment center.

[c] Further adjusted for education, Townsend deprivation index, and the number of clinic visits.

[d] Further adjusted for BMI, ever used alcohol, physical activity, diabetes, and stroke, further stratified by ever smoked.