Infectious disease burden and the risk of Alzheimer's disease: A population-based study

Authors: Antonios **Douros** MD, PhD[1,2,3,4], Christina **Santella** MSc[3], Sophie **Dell'Aniello** MSc[3], Laurent **Azoulay** PhD[2,3,5], Christel **Renoux** MD PhD[2,3,6], Samy **Suissa** PhD[1,2,3], Paul **Brassard** MD, MSc[1,2,3]

Affiliations

1. Department of Medicine, McGill University, Montreal, QC, Canada
2. Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada
3. Centre for Clinical Epidemiology, Lady Davis Institute, Montreal, QC, Canada
4. Institute of Clinical Pharmacology and Toxicology, Charité - Universitätsmedizin Berlin, Berlin, Germany
5. Gerald Bronfman Department of Oncology, McGill University, Montreal, QC, Canada
6. Department of Neurology and Neurosurgery, McGill University, Montreal, Québec, Canada

Corresponding author

Dr. Paul Brassard Centre for Clinical Epidemiology, Lady Davis Institute, Montreal, QC, Canada

3755 Chemin de la Cote St-Catherine, H-424, Montréal QC Canada H3T1E2

Tel 514 340 7563, paul.brassard@mcgill.ca

Abstract: 259 words

Manuscript: 3,540 words

February 23, 2021

The final publication is available at IOS Press through: http://dx.doi.org/10.3233/JAD-201534 .

This is an accepted manuscript of Douros, Antonios, Santella, Christina, Dell'Aniello, Sophie, Azoulay, Laurent, Renoux, Christel , Suissa, Samy, & Brassard, Paul. (2021). Infectious Disease Burden and the Risk of Alzheimer's Disease: A Population-Based Study. *Journal of Alzheimer's Disease*. doi:10.3233/JAD-201534 .

ABSTRACT

Background

Previous studies suggested a link between various infectious pathogens and the development of Alzheimer's disease (AD), posing the question whether infectious disease could present a novel modifiable risk factor.

Objective

To assess whether infectious disease burden due to clinically apparent infections is associated with an increased risk of AD.

Methods

We conducted a population-based nested case-control study using the United Kingdom Clinical Practice Research Datalink. We included all dementia-free subjects ≥50 years of age enrolling in the database between January 1988 and December 2017. Each case of AD identified during follow-up was matched with up to 40 controls. Conditional logistic regression estimated adjusted odds ratios (ORs) with 95% confidence intervals (CIs) of AD associated with ≥1 infection diagnosed >2 years before the index date compared with no infection during the study period. We further stratified by time since first infection and cumulative number of infections.

Results

The cohort included overall 4,262,092 individuals (mean age at cohort entry 60.4 years; 52% female). During a median follow-up of 10.5 years, 40,455 cases of AD were matched to 1,610,502 controls. Compared with having no burden of infectious disease, having a burden of infectious disease was associated with an increase in the risk of AD (OR, 1.05; 95% CI, 1.02 to 1.08). The

risk increased with longer time since first infection, peaking after 12-30 years (OR, 1.11; 95% CI, 1.05-1.17). The risk did not increase with cumulative number of infections.

Conclusion

The overall risk of AD associated with infectious disease burden was small but increased gradually with longer time since first infection.

KEYWORDS

Infection, dementia, Alzheimer's disease, neurodegenerative diseases, epidemiology

RUNNING TITLE

Infectious disease burden and Alzheimer's disease

INTRODUCTION

Dementia currently affects around 50 million people globally with nearly 10 million cases being newly diagnosed every year [1]. Alzheimer's disease is the most common form of dementia contributing up to 70% of cases [1]. Given the ageing population and increasing life expectancies, the burden of Alzheimer's disease is projected to dramatically increase in the following decades. Thus, ongoing research has been dedicated towards understanding the pathology of this disease in order to develop effective treatment and prevention strategies.

To date, several modifiable risk factors for Alzheimer's disease have been identified including smoking, obesity, or arterial hypertension [2]. However, randomized controlled trials studying the effects of multimodal interventions targeting several of these risk factors showed little [3] or no efficacy [4, 5]. Interestingly, evidence from many pre-clinical, serological, and post-mortem studies has suggested a link between various infectious pathogens and the development of Alzheimer's disease [6, 7, 8, 9, 10, 11, 12, 13], posing the question whether infectious disease could present a novel modifiable risk factor. Moreover, studies assessing the association between clinically apparent infections and the risk of Alzheimer's disease or overall dementia reported increased risks of up to 260% [14, 15, 16, 17, 18]. However, these studies had methodological limitations including reverse causality, selection bias, and important residual confounding, which render the interpretation of their findings difficult [14, 15, 16, 18]. In addition, the role of cumulative infectious disease burden and timing of infections with respect to Alzheimer's disease remains poorly understood.

Taken together, current literature lacks robust epidemiological evidence on the potential association between infectious disease burden and the risk of Alzheimer's disease. Thus, our

population-based nested case-control study assessed whether infectious disease burden, defined by clinically apparent infections easily detectable in routine clinical practice and related to pathogens previously linked to dementia, is associated with an increased risk of Alzheimer's disease.

METHODS

Data source

We conducted a population-based nested case-control study using the United Kingdom (UK) Clinical Practice Research Datalink (CPRD) Gold. The CPRD contains the medical records of over 11 million patients enrolled across 700 UK general practices and is one of the largest databases of longitudinal medical records from the primary care setting in the world [19]. Age, sex, and ethnicity distributions of patients in the CPRD are broadly representative of the UK population.[19] In addition, because general practitioners in the UK serve as a first point of contact for non-emergency health-related issues, the database contains useful information on routinely recorded symptoms, laboratory tests, diagnoses, therapies, health-related behaviors, and referrals to secondary care [19]. Medical diagnoses and procedures are recorded using the Read code classification, a hierarchical coding system containing over 80,000 terms encompassing the various aspects of a patient's health status [20]. The CPRD undergoes regular quality controls, and its valid and high-quality health data makes it a favorable data source for epidemiological research covering a vast range of health outcomes [19].

Study population

We included all subjects at least 50 years of age enrolled in the CPRD between January 1, 1988 and December 31, 2017. Cohort entry date was defined as the date of the 50th birthday of the subject or one year after their date of enrolment in the CPRD, whichever occurred later. We then excluded subjects with a prior diagnosis of any dementia, including mild cognitive impairment, and those with early symptoms suggestive of dementia (e.g., memory impairment, aphasia, apraxia,

or agnosia) at any time before cohort entry. We also excluded subjects treated with medications indicated for dementia including acetylcholinesterase inhibitors (i.e., donepezil, rivastigmine, or galantamine) and N-methyl-D-aspartate receptor antagonists (i.e., memantine) at any time before cohort entry. Cohort members were followed from the date of cohort entry until the date of the first outcome event (defined below), end of registration with the general practice, death from any cause, or the end of the study period (i.e., 31 December 2019), whichever occurred first.

Case definition

Within the study cohort, we identified all subjects with a first-ever diagnosis of Alzheimer's disease at any time after cohort entry. We defined Alzheimer's disease based on a modified algorithm initially developed and validated by Imfeld and colleagues [21], which has previously been used by our group [22]. Using this algorithm, Alzheimer's disease was defined by meeting at least one of the following criteria: (i) a diagnosis of Alzheimer's disease with at least one prescription of a medication for dementia, (ii) a diagnosis of unspecified dementia with at least two prescriptions of a medication for dementia, (iii) at least two diagnoses of Alzheimer's disease, (iv) a diagnosis of Alzheimer's disease after a dementia test (e.g., Mini-Mental State Examination, abbreviated mental test) or a referral to a specialist (e.g., neurologist, psychiatrist, geriatrician, psychogeriatrician) or a neuroimaging assessment (e.g., magnetic resonance imaging, computed tomography, single-photon emission computed tomography), or (v) a diagnosis of Alzheimer's disease with any dementia symptoms (e.g., memory impairment, aphasia, apraxia, agnosia) in any sequence. The index date (i.e., date of Alzheimer's disease diagnosis) was defined as the date of

the last event contributing to the definition. The quality of the recording of Alzheimer's disease in the CPRD has been shown to be high, with a positive predictive value of 83% [23].

Control selection

Each case of Alzheimer's disease was matched with up to 40 controls who belonged to the risk set defined by the case (i.e., those subjects still at risk of the event at the index date) on age (±1 year), sex, cohort entry date (±1 year), and duration of follow-up. The high number of controls was chosen to minimize feasibility issues in secondary analyses related to the potential scarcity of matched controls. The date resulting in the same duration of follow-up for the case and controls was set as the index date for the controls. Controls could contribute to different risk sets and could subsequently become a case. For our analyses, we only used cases and controls with at least two years follow-up given the use of a two-year lag period in the assessment of exposure (see below).

Exposure definition

For cases and controls with at least two years of follow-up, we identified all diagnoses of clinically apparent infections potentially involving pathogens which have previously been linked to the pathophysiology of Alzheimer's disease regardless of the proposed mechanism. These infections included herpes labialis or genitalis (*Herpes simplex virus*) [9], cytomegalovirus related hepatitis, retinitis, colitis, mononucleosis, or other infections [10], Lyme disease (*Borrelia burgdorferi*) [11, 24], gingivitis (*Porphyromonas gingivalis*) [12], urinary tract infections

(*Escherichia coli*) [13], gastritis (*Helicobacter pylori*) [8], pneumonia (*Chlamydophila pneumonia)* [7, 25], and candidiasis (*Candida albicans*) [6]. Infections due to pathogens with no potential link to the pathophysiology of Alzheimer's disease (e.g., influenza, common cold) were not considered in the analyses. Subjects with a clinical diagnosis of any of these infections two years or more before the index date were considered as having a burden of infectious disease, while those without a diagnosis of any of these infections during that time period were considered as having no burden of infectious disease. Subjects with a diagnosis of any of these infections only within the 2-year period before the index date were also considered as having no burden of infectious disease. This 2-year 'lag period' was introduced given the insidious (i.e., non-acute) nature of the study outcome, and also to account for the delays associated with the diagnosis of Alzheimer's disease [2, 26].

Statistical analysis

Conditional logistic regression was used to compute odds ratios of Alzheimer's disease associated with infectious disease burden, compared with no infectious disease burden. Odds ratios are unbiased estimators of hazard ratios, with little or no loss in precision [27, 28]. In addition to the matching factors, estimates were further adjusted in the regression model for the following potential confounders associated with Alzheimer's disease, measured at any time before cohort entry: body mass index category (<25 kg/m^2, 25-29 kg/m^2, ≥ 30 kg/m^2, unknown; last measurement before cohort entry), smoking status (ever, never, unknown), alcohol-related disorders (including alcoholism, alcoholic cirrhosis, or alcoholic hepatitis), arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression,

epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, and cancer. We also included the use of the following drugs in the two years prior to the index date: oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants. In the case of missing data (expected for the covariates body mass index and smoking), a separate category ('unknown') was created to classify this missing information.

Secondary analyses

We conducted five exploratory secondary analyses. First, to examine a potential 'dose-response' relation between infectious disease burden and the risk of Alzheimer's disease, we estimated odds ratios for each of the following categories: 1, 2-3, and >3 infections. Second, to examine a potential 'time-response' relation between infectious disease burden and the risk of Alzheimer's disease, we estimated odds ratios for each of the following categories: 0-4.9, 5-7.9, 8-11.9, and 12-30 years since the time of the first infection (first infection after the 50th birthday; cut-offs for the different categories were based on the distribution of durations of follow-up among the controls). To account for the scenario of a non-linear association, we also modeled time since first infection as a continuous variable using restricted cubic splines with five interior knots [29]. Third, we examined the association by specific type of infection (i.e., herpes, cytomegalovirus related infection, Lyme disease, gingivitis, urinary tract infection, gastritis, pneumonia, candidiasis). Finally, we stratified by age (<65 years versus ≥65 years) and sex to assess a potential effect

modification, since advanced age and female sex are established risk factors of Alzheimer's disease [30, 31].

Sensitivity analyses

We also performed several sensitivity analyses to assess the robustness of our findings. First, given the uncertainty regarding the latency of a potential association between infectious disease burden and the development of Alzheimer's disease, we repeated the primary analysis after increasing the lag period to 3, 5, and 10 years. Second, we censored follow-up at dementia diagnoses of non-Alzheimer's disease etiology (e.g., vascular dementia, alcoholic dementia). Third, we restricted the medical codes for pneumonia to those with a clear link to *Chlamydophila pneumonia* to reduce exposure misclassification due to pneumonia caused by other infectious pathogens (e.g., pneumococci, viruses; medical codes for other infections remained unchanged). Finally, given that some of the previous studies assessed the association between infectious disease burden and the risk of overall dementia (instead of Alzheimer's disease specifically) [16, 17], we repeated the analyses after expanding our outcome definition to include any diagnosis of dementia (see **Supplementary Table 1** for Read codes). All analyses were conducted with SAS version 9.4 (SAS institute, Cary, NC).

Standard Protocol Approvals, Registrations, and Patient Consents

The study protocol was approved by the Independent Scientific Advisory Committee of the CPRD (protocol 19_236R) and by the Research Ethics Board of the Jewish General Hospital,

Montreal, Canada. Written consent from participants was not required due to use of anonymized data and no direct patient involvement.

RESULTS

The study cohort included a total of 4,262,092 dementia free individuals who were at least 50 years of age and enrolled in the CPRD between January 1, 1988 and December 31, 2017 and followed until December 31, 2019 (**Figure 1**). Mean (standard deviation) age at cohort entry was 60.4 (11.5) years, and 52.1% were female. During a median (interquartile range) follow-up of 10.5 (6.2 to 14.6) years, 42,912 individuals in the study cohort were diagnosed with Alzheimer's disease (crude incidence rate, 2.3 per 1000 person-years). Most diagnoses were based on the combination of a diagnostic code of Alzheimer's disease accompanied either by respective symptoms, or tests for dementia, referrals to specialists, and neuroimaging assessments (**Supplementary Table 2**).

We matched 40,455 cases of Alzheimer's disease with at least two years of follow-up to 1,610,502 controls from the study cohort. Characteristics of cases and their matched controls are presented in **Table 1**. Cases were similar to controls except that they were more likely to have previously used antipsychotics or antidepressants. Compared with having no burden of infectious disease, having a burden of infectious disease was associated with a small increase in the risk of Alzheimer's disease (odds ratio, 1.05; 95% confidence interval, 1.02 to 1.08) (**Table 2**). There was no evidence of a dose-response relation, with the risk of Alzheimer's disease not significantly changing with cumulative number of infections (**Table 2**). However, there was a suggestion of a time-response relation, with the risk of Alzheimer's disease gradually increasing with longer time intervals since the first infection (peak after 12-30 years: odds ratio, 1.11; 95% confidence interval, 1.05 to 1.17; p for trend = 0.0003) (**Table 2**, **Figure 2**).

Stratifying by specific type showed an increased risk for gastritis (odds ratio, 1.08; 95% confidence interval, 1.03 to 1.13) but not for other infections (**Table 3**). Age did not seem to modify

the association; however, the risk of Alzheimer's disease was only increased among female patients (odds ratio, 1.08; 95% confidence interval, 1.04 to 1.11) and not among male patients (odds ratio, 0.99; 95% confidence interval, 0.94 to 1.04) (**Supplementary Table 3**).

Finally, the sensitivity analyses using extended lag periods, censoring follow-up at non-Alzheimer's disease dementia diagnoses, and restricting pneumonia diagnoses to those with a clear link to *Chlamydophila* yielded results that were highly consistent with those of the primary analysis (**Supplementary Table 4**). The results also did not change substantially after expanding our outcome definition to include any dementia (characteristics of cases of dementia and their matched controls are presented in **Supplementary Table 5**; the results of the primary, secondary, and sensitivity analyses are presented in **Supplementary Tables 6-9** and **Supplementary Figure 1**). For example, similar to the analyses on the risk of Alzheimer's disease, the increased risk of any dementia associated with infectious disease burden was not accompanied by a dose-response relation but a possible time-response relation. However, there was an increased risk associated with pneumonia, which was not observed in the Alzheimer's disease specific analyses.

DISCUSSION

Our large population-based nested case-control study showed a small increase in the risk of Alzheimer's disease associated with infectious disease burden. This effect was not augmented with cumulative number of infections, but there was a suggestion of a gradual increase in the risk with longer time since the first infection. Focusing on specific types of infections, we identified a small increase in the risk associated with gastritis. Moreover, sex seemed to modify the association, with the risk of Alzheimer's disease being increased only among female patients. The results remained consistent in sensitivity analyses addressing different sources of bias.

Despite the rapidly increasing numbers of individuals diagnosed with Alzheimer's disease and the devastating course of the disease, the efficacy of available pharmacologic treatments is modest at best [2]. Moreover, multimodal interventions targeting several modifiable risk factors of Alzheimer's disease and dementia have yielded sobering findings [3, 4, 5]. As a result, there is an ongoing search for novel angles in the area of Alzheimer's disease prevention, with one of the most promising approaches in the past years being the 'infectious hypothesis' [32]. According to this hypothesis, hallmarks of Alzheimer's disease such as the deposition of amyloid-β peptide or abnormal forms of tau protein in the brain are indicators of an infectious etiology [32]. Of note, these pathological changes may occur up to 20 years prior to the onset of symptoms [33]. The obvious and extremely intriguing consequence, should the infectious hypothesis be proven, would be that by reducing the burden of infectious diseases (e.g., via preventive treatments or vaccination programs) we could also potentially reduce the burden of Alzheimer's disease.

Several pre-clinical, serological, and post-mortem studies have supported this hypothesis linking various infectious pathogens to Alzheimer's disease [6, 7, 8, 9, 10, 11, 12, 13]. Moreover,

epidemiological studies have uniformly shown an increased risk of Alzheimer's disease or overall dementia associated with clinically apparent infections (e.g., pneumonia, septicemia, gingivitis, or overall infections), which ranged from 20% up to 260% [14, 15, 16, 17, 18]. However, the quality of these studies could be affected by reverse causality [15, 16, 18], selection bias [14, 17], and important residual confounding [15, 16, 18]. Reverse causality in particular can lead to spuriously increased effect estimates in this setting, since patients with early symptoms of Alzheimer's disease could be at a higher risk of infections, or they could be followed-up more closely by the treating physician increasing the probability of infectious disease reporting [2]. Of note, the study with the highest quality included almost exclusively male individuals, which could compromise external validity [17].

Our study also showed a statistically significant increase in the risk of Alzheimer's disease associated with infectious disease burden as defined by clinically apparent infections. Of note, the increase (5%) was much smaller than what has previously been reported, which potentially limits the clinical significance of the association. Moreover, there was no further increase in the risk of Alzheimer's disease with cumulative number of infections. However, there was a gradual increase in the risk with longer time since the first infection, with a peak (11%) after 12-30 years. The potential time-response relation is intriguing, suggesting that infections occurring many years before the diagnosis of Alzheimer's disease may contribute to its etiology. This hypothesis is in accordance with the early, pre-symptomatic onset of pathological changes linked to infections such as amyloid-β peptide deposition or tau protein abnormalities discussed earlier. That being said, additional studies are needed in the area to better understand this potential association.

After stratifying by sex, we observed an increased risk of Alzheimer's disease associated with infectious disease burden only among female patients, a finding that supports previously reported data on the effect modifying properties of female sex [34]. When focusing on specific types of infection, we observed a potential signal for gastritis but not for other entities. Of note, while these analyses were pre-specified and based on a sufficient number of exposed cases, their findings should be considered hypothesis generating given the amount of assessed associations. Thus, they require further investigation. Finally, another finding warranting additional research is the increased risk of any dementia associated with pneumonia, which was not observed in the Alzheimer's disease specific analysis.

Our study has several strengths. First, the population-based design and the application of few exclusion criteria during the construction of the study cohort likely maximized the generalizability of our findings. Second, the large sample size allowed the calculation of precise effect estimates in the primary analysis and the secondary analyses. Indeed, the secondary analyses assessing potential dose-response and time-response relations between infectious disease burden and the risk of Alzheimer's disease yielded useful insight regarding aspects of the association that were poorly characterized so far. Finally, the use of a 2-year lag period (and even longer lag periods in sensitivity analyses) minimized the possibility of reverse causality, a well-established challenge when assessing insidious, non-acute outcomes such as Alzheimer's disease [2].

This study has some limitations. First, given its observational nature, residual confounding is possible. To mitigate this potential limitation, we matched on age and sex and further adjusted for numerous important confounders. Second, misclassification of exposure is possible, since we did not have access to microbiology data to confirm the infection. For example, not every gastritis

case is a result of infection, with medications such as non-steroidal anti-inflammatory drugs or stress being possible alternative causes. Moreover, the link between the infectious pathogen with the putative role in the pathophysiology of Alzheimer's disease and the clinically apparent infection may be weak. For example, pneumococci and viruses are far more common causes of pneumonia than *Chlamydophila pneumonia*. However, a sensitivity analysis restricting to pneumonias related to *Chlamydophila pneumonia* yielded highly consistent results. Moreover, we would like to emphasize that the goal of our study was to specifically focus on infections that are symptomatic and thus easily detectable in the natural setting of routine clinical practice. Third, we assessed the infectious disease burden only in patients at least 50 years of age. Thus, infections occurring earlier in life could not be considered in our analyses. Given the observed time-response relation between time since first infection and the risk of Alzheimer's disease, future studies should assess the potential impact of infectious disease burden in the first decades of adulthood. Fourth, misclassification of the outcome is possible. However, the recording of Alzheimer's disease and dementia in general in the CPRD has been shown to be good [35]. Moreover, we defined Alzheimer's disease using a previously validated algorithm which incorporates not only diagnostic codes but also symptoms, diagnostic procedures, and medications, which possibly further improved the accuracy of our outcome definition [21]. In addition, the incidence rate in our study (2.3 per 1000 person-years) was consistent with the incidence rates reported in other population based studies with similar age distributions (from 1.7 per 1000 person-years to 7.1 per 1000 person-years for individuals aged between 65 and 75 years) [36, 37]. Finally, since we did not have access to patients' vitamin D levels, analyses considering the potential role of vitamin D as a risk factor of infection-associated Alzheimer's disease were not possible.

Overall, our large population-based nested case-control study identified a statistically significant but small and probably not clinically significant increase in the risk of Alzheimer's disease associated with infectious disease burden. Given that the risk seemed to gradually increase with longer time since the first infection, peaking after 12 years, the role of infections occurring several years prior to the diagnosis of Alzheimer's disease warrants further investigation.

ACKNOWLEDGMENTS

This study was funded by a Discovery Proof of Concept Grant of the Alzheimer Society of Canada Research Program (Grant number: 21-13) to P. Brassard. A. Douros and P. Brassard had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. A. Douros, C. Santella and P. Brassard designed the study. C. Santella, S. Dell'Aniello, A. Douros, S. Suissa and P. Brassard contributed to the data analysis. A. Douros wrote the initial version of the manuscript. L. Azoulay, C. Renoux, S. Dell'Aniello and S. Suissa reviewed the manuscript for content, interpretation, and accuracy. All authors reviewed and accepted the final version for submission. A. Douros is the recipient of Chercheur-Boursier Junior 1 Award from the *Fonds de recherche du Québec – santé* (FRQS). C. Renoux is the recipient of Chercheur-Boursier Junior 2 Award from the FRQS. L. Azoulay is the recipient of Chercheur-Boursier Senior Award from the FRQS and a William Dawson Scholar award from McGill University.

CONFLICT OF INTEREST/DISCLOSURE STATEMENT

L. Azoulay served as a consultant for Janssen and Pfizer for work unrelated to this study. All other authors declare no conflict of interest. None of the authors have published, posted, or submitted any related papers from the same study.

REFERENCES

[1] World Health Organization (2019) Dementia. https://www.who.int/newsroom/fact-sheets/detail/dementia, Accessed on May 20, 2020.

[2] Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, Ballard C, Banerjee S, Burns A, Cohen-Mansfield J, Cooper C, Fox N, Gitlin LN, Howard R, Kales HC, Larson EB, Ritchie K, Rockwood K, Sampson EL, Samus Q, Schneider LS, Selbæk G, Teri L, Mukadam N (2017) Dementia prevention, intervention, and care. *Lancet* **390**, 2673-734.

[3] Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, Bäckman L, Hänninen T, Jula A, Laatikainen T, Lindström J, Mangialasche F, Paajanen T, Pajala S, Peltonen M, Rauramaa R, Stigsdotter-Neely A, Strandberg T, Tuomilehto J, Soininen H, Kivipelto M (2015) A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet* **385**, 2255-63.

[4] Andrieu S, Guyonnet S, Coley N, Cantet C, Bonnefoy M, Bordes S, Bories L, Cufi MN, Dantoine T, Dartigues JF, Desclaux F, Gabelle A, Gasnier Y, Pesce A, Sudres K, Touchon J, Robert P, Rouaud O, Legrand P, Payoux P, Caubere JP, Weiner M, Carrié I, Ousset PJ, Vellas B; MAPT Study Group (2017) Effect of long-term omega 3 polyunsaturated fatty acid supplementation with or without multidomain intervention on cognitive function in elderly adults with memory complaints (MAPT): a randomised, placebo-controlled trial. *Lancet Neurol* **16**, 377-89.

[5] Moll van Charante EP, Richard E, Eurelings LS, van Dalen JW, Ligthart SA, van Bussel EF, Hoevenaar-Blom MP, Vermeulen M, van Gool WA (2016) Effectiveness of a 6-year multidomain vascular care intervention to prevent dementia (preDIVA): a cluster-randomised controlled trial. *Lancet* **388,** 797-805.

[6] Alonso R, Pisa D, Aguado B, Carrasco L (2017) Identification of Fungal Species in Brain Tissue from Alzheimer's Disease by Next-Generation Sequencing. *J Alzheimers Dis* **58**, 55-67.

[7] Balin BJ, Little CS, Hammond CJ, Appelt DM, Whittum-Hudson JA, Gérard HC, Hudson AP (2008) Chlamydophila pneumoniae and the etiology of late-onset Alzheimer's disease. *J Alzheimers Dis* **13**, 371-80.

[8] Beydoun MA, Beydoun HA, Elbejjani M, Dore GA, Zonderman AB (2018) Helicobacter pylori seropositivity and its association with incident all-cause and Alzheimer's disease dementia in large national surveys. *Alzheimers Dement* **14**, 1148-58.

[9] Itzhaki RF (2017) Herpes simplex virus type 1 and Alzheimer's disease: possible mechanisms and signposts. *FASEB J* **31**, 3216-26.

[10] Lövheim H, Olsson J, Weidung B, Johansson A, Eriksson S, Hallmans G, Elgh F (2018) Interaction between Cytomegalovirus and Herpes Simplex Virus Type 1 Associated with the Risk of Alzheimer's Disease Development. *J Alzheimers Dis* **61**, 939-45.

[11] Miklossy J (2011) Alzheimer's disease - a neurospirochetosis. Analysis of the evidence following Koch's and Hill's criteria. *J Neuroinflammation* **8**, 90.

[12] Singhrao SK, Harding A, Poole S, Kesavalu L, Crean S (2015) Porphyromonas gingivalis Periodontal Infection and Its Putative Links with Alzheimer's Disease. *Mediators Inflamm* **2015**, 137357.

[13] Zhan X, Stamova B, Jin LW, DeCarli C, Phinney B, Sharp FR (2016) Gram-negative bacterial molecules associate with Alzheimer disease pathology. *Neurology* **87**, 2324-32.
[14] Chen CK, Wu YT, Chang YC (2017) Association between chronic periodontitis and the risk of Alzheimer's disease: a retrospective, population-based, matched-cohort study. *Alzheimers Res Ther* **9**, 56-56.
[15] Chou CH, Lee JT, Lin CC, Sung YF, Lin CC, Muo CH, Yang FC, Wen CP, Wang IK, Kao CH, Hsu CY, Tseng CH (2017) Septicemia is associated with increased risk for dementia: a population-based longitudinal study. *Oncotarget* **8**, 84300-08.
[16] Dunn N, Mullee M, Perry VH, Holmes C (2005) Association between dementia and infectious disease: evidence from a case-control study. *Alzheimer Dis Assoc Disord* **19**, 91-94.
[17] Mawanda F, Wallace RB, McCoy K, Abrams TE (2016) Systemic and localized extra-central nervous system bacterial infections and the risk of dementia among US veterans: A retrospective cohort study. *Alzheimers Dement* **4**, 109-17.
[18] Tate JA, Snitz BE, Alvarez KA, Nahin RL, Weissfeld LA, Lopez O, Angus DC, Shah F, Ives DG, Fitzpatrick AL, Williamson JD, Arnold AM, DeKosky ST, Yende S; GEM Study Investigators (2014) Infection hospitalization increases risk of dementia in the elderly. *Crit Care Med* **42**, 1037-46.
[19] Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L (2015) Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* **44**, 827-36.
[20] Chisholm J (1990) The Read clinical classification. *BMJ* **300**,1092.
[21] Imfeld P, Brauchli Pernus YB, Jick SS, Meier CR (2013) Epidemiology, co-morbidities, and medication use of patients with Alzheimer's disease or vascular dementia in the UK. *J Alzheimers Dis* **35**, 565-73.
[22] Sinyavskaya L, Gauthier S, Renoux C, Dell'Aniello S, Suissa S, Brassard P (2018) Comparative effect of statins on the risk of incident Alzheimer disease. *Neurology* **90**, e179-e87.
[23] Khan NF, Harrison SE, Rose PW (2010) Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* **60**, e128-36.
[24] Miklossy J, Kis A, Radenovic A, Miller L, Forro L, Martins R, Reiss K, Darbinian N, Darekar P, Mihaly L, Khalili K (2006) Beta-amyloid deposition and Alzheimer's type changes induced by Borrelia spirochetes. *Neurobiol Aging* **27**, 228-36.
[25] Little CS, Hammond CJ, MacIntyre A, Balin BJ, Appelt DM (2004) Chlamydia pneumoniae induces Alzheimer-like amyloid plaques in brains of BALB/c mice. *Neurobiol Aging* **25**, 419-29.
[26] Rothman KJ (1981) Induction and latent periods. *Am J Epidemiol* **114**, 253-9.
[27] Essebag V, Platt RW, Abrahamowicz M, Pilote L (2005) Comparison of nested case-control and survival analysis methodologies for analysis of time-dependent exposure. *BMC Med Res Methodol* **5,** 5.
[28] Suissa S (2006) Novel Approaches to Pharmacoepidemiology Study Design and Statistical Analysis. *In Pharmacoepidemiology*, B.L. Strom, ed. pp. 811-29.
[29] Durrleman S, Simon R (1989) Flexible regression models with cubic splines. *Stat Med* **8**, 551-61.
[30] van der Flier WM, Scheltens P (2005) Epidemiology and risk factors of dementia. *J Neurol Neurosurg Psychiatry* **76** Suppl 5, v2-7.

[31] Seshadri S, Wolf PA, Beiser A, Au R, McNulty K, White R, D'Agostino RB (1997) Lifetime risk of dementia and Alzheimer's disease. The impact of mortality on risk estimates in the Framingham Study. *Neurology* **49**, 1498-504.
[32] Itzhaki RF, Lathe R, Balin BJ, Ball MJ, Bearer EL, Braak H, Bullido MJ, Carter C, Clerici M, Cosby SL, Del Tredici K, Field H, Fulop T, Grassi C, Griffin WS, Haas J, Hudson AP, Kamer AR, Kell DB, Licastro F, Letenneur L, Lövheim H, Mancuso R, Miklossy J, Otth C, Palamara AT, Perry G, Preston C, Pretorius E, Strandberg T, Tabet N, Taylor-Robinson SD, Whittum-Hudson JA (2016) Microbes and Alzheimer's Disease. *J Alzheimers Dis* **51**, 979-84.
[33] Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, Bakardjian H, Benali H, Bertram L, Blennow K, Broich K, Cavedo E, Crutch S, Dartigues JF, Duyckaerts C, Epelbaum S, Frisoni GB, Gauthier S, Genthon R, Gouw AA, Habert MO, Holtzman DM, Kivipelto M, Lista S, Molinuevo JL, O'Bryant SE, Rabinovici GD, Rowe C, Salloway S, Schneider LS, Sperling R, Teichmann M, Carrillo MC, Cummings J, Jack CR Jr; Proceedings of the Meeting of the International Working Group (IWG) and the American Alzheimer's Association on "The Preclinical State of AD"; July 23, 2015; Washington DC, USA (2016) Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimers Dement* **12**, 292-323.
[34] Rahman A, Jackson H, Hristov H, Isaacson RS, Saif N, Shetty T, Etingin O, Henchcliffe C, Brinton RD, Mosconi L (2019) Sex and Gender Driven Modifiers of Alzheimer's: The Role for Estrogenic Control Across Age, Race, Medical, and Lifestyle Risks. *Front Aging Neurosci* **11**, 315.
[35] McGuinness LA, Warren-Gash C, Moorhouse LR, Thomas SL (2019) The validity of dementia diagnoses in routinely collected electronic health records in the United Kingdom: A systematic review. *Pharmacoepidemiol Drug Saf* **28**, 244-255.
[36] Brookmeyer R, Gray S, Kawas C (1998) Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health* **88**, 1337-42.
[37] Rajan KB, Weuve J, Barnes LL, Wilson RS, Evans DA (2019) Prevalence and incidence of clinically diagnosed Alzheimer's disease dementia from 1994 to 2012 in a population study. *Alzheimers Dement* **15**, 1-7.

FIGURE LEGENDS

Figure 1. Flowchart showing the construction of the study cohort

Abbreviations: CPRD, Clinical Practice Research Datalink.

Figure 2. Restricted cubic spline of time since first infection on the risk of Alzheimer's disease

The solid line shows the odds ratio and the dashed lines show the lower and upper bound of the 95% confidence interval. The curve begins at 2 years given the use of a 2-year lag period in the definition of exposure.

Figure 1

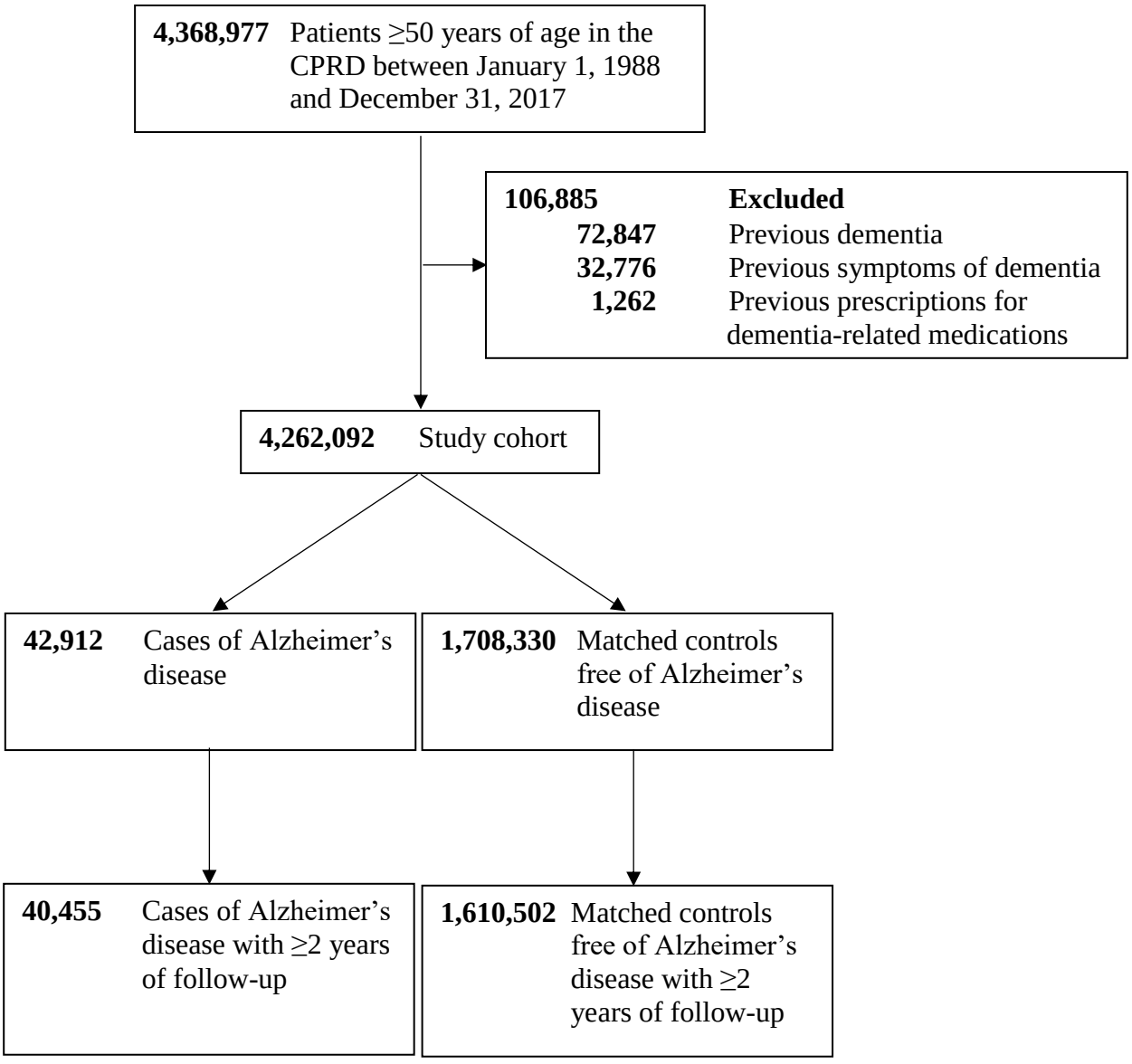

4,368,977 Patients ≥50 years of age in the CPRD between January 1, 1988 and December 31, 2017
106,885 Excluded
72,847 Previous dementia
32,776 Previous symptoms of dementia
1,262 Previous prescriptions for dementia-related medications
4,262,092 Study cohort
42,912 Cases of Alzheimer's disease
1,708,330 Matched controls free of Alzheimer's disease
40,455 Cases of Alzheimer's disease with ≥2 years of follow-up
1,610,502 Matched controls free of Alzheimer's disease with ≥2 years of follow-up

Figure 2

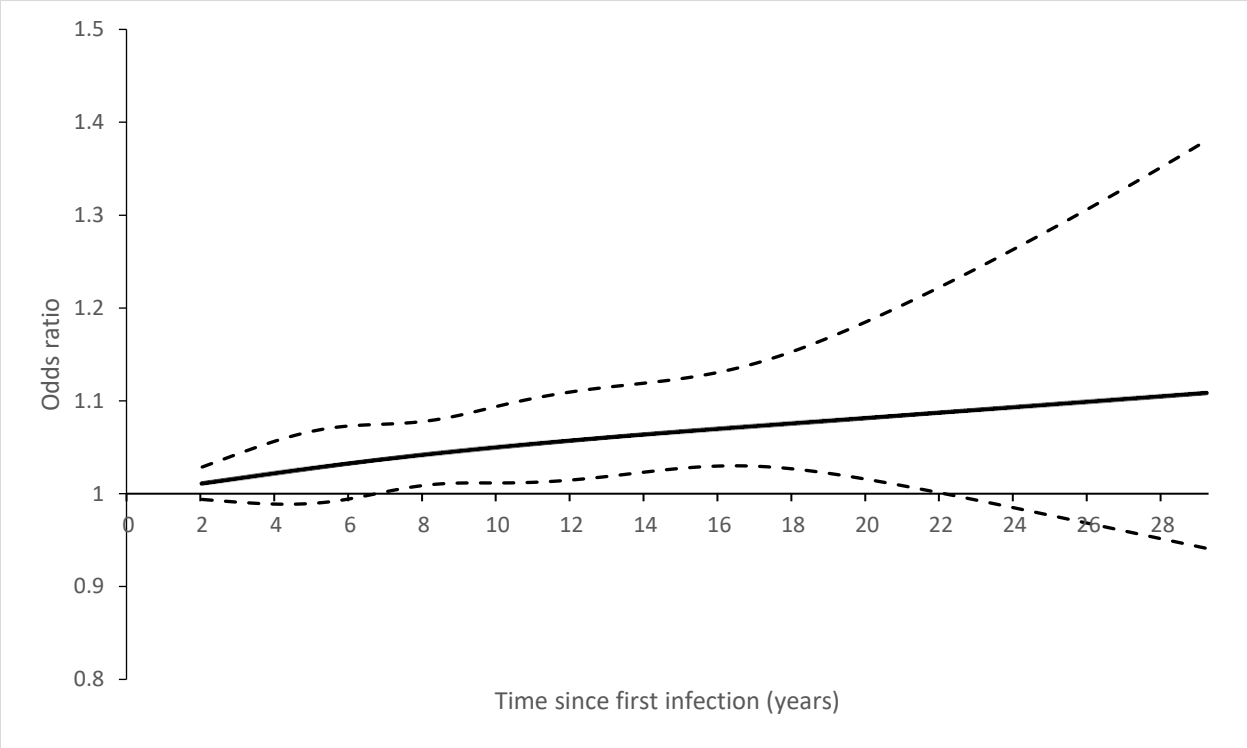

1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
Odds ratio
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
Time since first infection (years)

Table 1. Baseline characteristics of cases of Alzheimer's disease and their matched controls[*]

Characteristic	Cases (n = 40,455)	Controls [a] (n = 1,610,502)
Age in years, mean (standard deviation)	80.6 (7.6)	80.6 (7.6)
Follow-up in years, mean (standard deviation)	11.3 (5.5)	11.3 (5.5)
Male sex, n (%)	14,454 (35.7)	573,799 (35.7)
Body mass index in kg/m^2		
< 25	11,499 (28.4)	403,345 (25.0)
25-29	10,191 (25.2)	402,459 (25.0)
≥ 30	4,331 (10.7)	192,114 (11.9)
Unknown	14,434 (35.7)	612,584 (38.1)
Smoking		
Ever	12,887 (31.9)	491,854 (30.5)
Never	17,794 (44.0)	694,250 (43.1)
Unknown	9,774 (24.2)	424,398 (26.4)
Alcohol-related disorders	649 (1.6)	23,301 (1.4)
Arterial hypertension	12,259 (30.3)	512,105 (31.8)
Atrial fibrillation	970 (2.4)	44,075 (2.7)
Congestive heart failure	686 (1.7)	30,296 (1.9)
Coronary artery disease	5,753 (14.2)	226,400 (14.1)
Stroke or transient ischemic attack	1,264 (3.1)	60,708 (3.8)
Peripheral vascular disease	824 (2.0)	32,873 (2.0)
Dyslipidemia	4,540 (11.2)	160,584 (10.0)
Diabetes mellitus	3,011 (7.4)	114,001 (7.1)
Chronic kidney disease	856 (2.1)	35,296 (2.2)
Liver disease	128 (0.3)	5,383 (0.3)
Depression	5,136 (12.7)	174,681 (10.8)
Epilepsy	583 (1.4)	19,623 (1.2)
Parkinson's disease	155 (0.4)	5,161 (0.3)
Previous traumatic brain injury	S	32 (0.0)
Osteoporosis	1,371 (3.4)	47,716 (3.0)
Hypothyroidism	3,461 (8.6)	135,212 (8.4)
Cancer	3,479 (8.6)	134,659 (8.4)
Medications [b]		
Oral anticoagulants	3,316 (8.2)	151,295 (9.4)
Antiplatelet agents	15,877 (39.3)	576,289 (35.8)
Opioids	15,855 (39.2)	634,538 (39.4)
Lipid-lowering drugs	17,409 (43.0)	665,481 (41.3)
Antihypertensives [c]	25,720 (63.6)	1,081,376 (67.2)
Antipsychotics	5,580 (13.8)	125,928 (7.8)

Non-steroidal anti-inflammatory drugs	7,217 (17.8)	320,862 (19.8)
Antidepressants	14,127 (34.9)	316,909 (19.7)

S = Cells with less than 5 counts are suppressed as per the confidentiality policies of the Clinical Practice Research Datalink.

[*] Numbers are presented as n (%) unless otherise specified.

[a] For controls, means and percentages were weighted by the inverse number of controls matched to each case.

[b] Measured in the two years prior to index date as a surrogate measure of overall health.

[c] Includes beta-blockers, thiazide diuretics, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, and calcium channel blockers.

Table 2. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of Alzheimer's disease (overall and stratified by cumulative number of infections and time since first infection)

Exposure	Cases (n = 40,455) n (%)	Controls (n = 1,610,502) n (%)	Crude** OR (95% CI)	Adjusted*** OR (95% CI)
Primary analysis				
No infections	32,405 (80.1)	1,321,474 (82.0)	Reference	Reference
≥1 infection	8,050 (19.9)	289,028 (18.0)	1.15 (1.12 to 1.18)	1.05 (1.02 to 1.08)
Cumulative number of infections ('dose-response')				
No infections	32,405 (80.1)	1,321,474 (82.0)	Reference	Reference
1	5,165 (12.8)	184,499 (11.5)	1.15 (1.12 to 1.18)	1.07 (1.04 to 1.11) ±
2-3	2,089 (5.2)	75,521 (4.7)	1.14 (1.09 to 1.19)	1.02 (0.97 to 1.06) ±
>3	796 (2.0)	29,008 (1.8)	1.13 (1.06 to 1.22)	0.97 (0.90 to 1.04) ±
Time since first infection in years ('time-response')*				
No infections	32,405 (80.1)	1,321,474 (82.0)	Reference	Reference
2-4.9	2,422 (6.0)	86,116 (5.4)	1.15 (1.10 to 1.20)	1.06 (1.01 to 1.10) ±±
5-7.9	1,840 (4.6)	69,337 (4.3)	1.09 (1.04 to 1.14)	1.00 (0.95 to 1.05) ±±
8-11.9	1,811 (4.5)	66,166 (4.1)	1.13 (1.08 to 1.19)	1.03 (0.98 to 1.08) ±±
12-30	1,977 (4.9)	67,409 (4.2)	1.22 (1.16 to 1.29)	1.11 (1.05 to 1.17) ±±

Abbreviations: OR, odds ratio; CI, confidence interval.

* Given the use of a 2-year lag period in the definition of exposure, the minimum time since first infection was 2 years.
** Matched on age, sex, date of cohort entry and duration of follow-up.
*** Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral

anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.
[±] P value for trend was 0.13 in the dose-response analysis
[±±]P value for trend was 0.0003 in the time-response analysis.

Table 3. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of Alzheimer's disease (stratified by specific type of infection)

Exposure	Cases (n = 40,455) n (%)	Controls (n = 1,610,502) n (%)	Crude** OR (95% CI)	Adjusted*** OR (95% CI)
Specific type*¥				
No infections	32,405 (80.1)	1,321,474 (82.0)	Reference	Reference
Urinary tract infections	5,807 (14.4)	210,896 (13.1)	1.12 (1.09 to 1.15)	1.03 (1.00 to 1.06)
Herpes	83 (0.2)	2,670 (0.2)	1.24 (0.99 to 1.54)	1.15 (0.92 to 1.43)
Lyme disease	S	200 (0.01)	0.60 (0.19 to 1.88)	0.60 (0.19 to 1.87)
Gingivitis	228 (0.6)	7,683 (0.5)	1.18 (1.03 to 1.35)	1.06 (0.93 to 1.21)
Gastritis	1,715 (4.2)	57,049 (3.5)	1.21 (1.15 to 1.27)	1.08 (1.03 to 1.13)
Pneumonia	481 (1.2)	19,185 (1.2)	0.99 (0.90 to 1.09)	0.92 (0.84 to 1.01)
Candidiasis	885 (2.2)	29,360 (1.8)	1.21 (1.13 to 1.29)	1.07 (1.00 to 1.15)

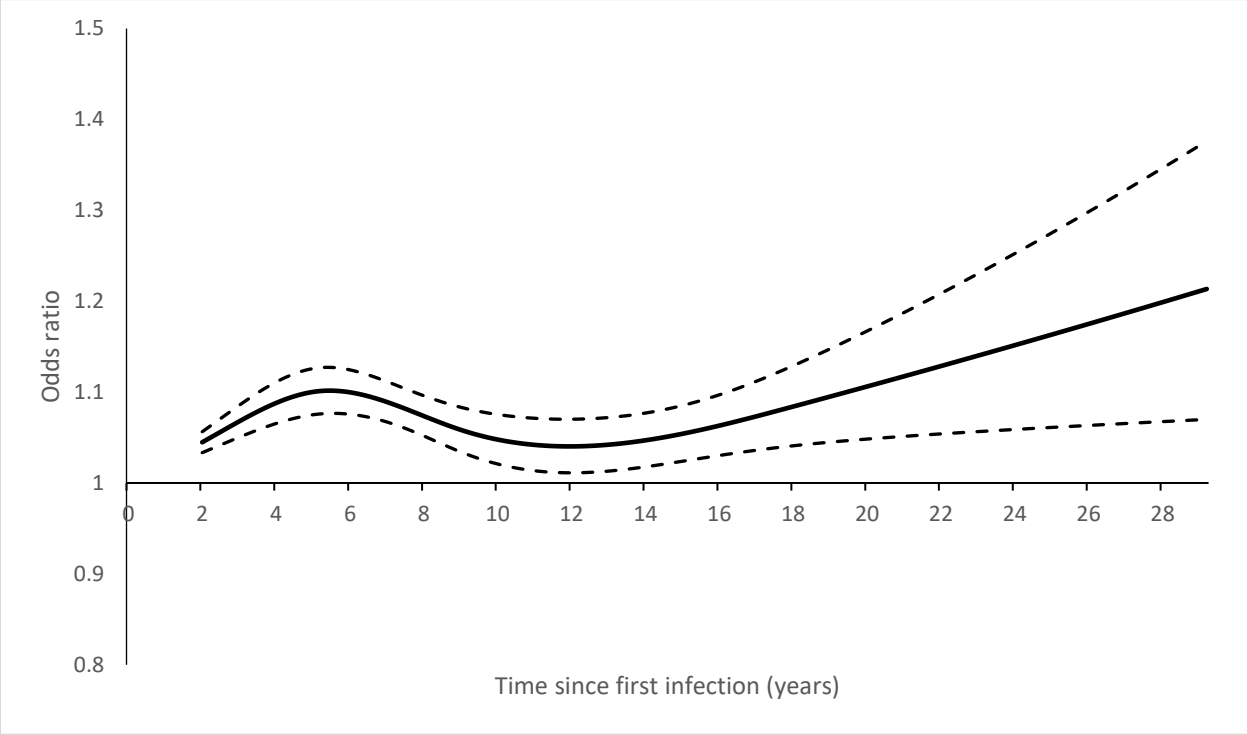
Abbreviations: OR, odds ratio; CI, confidence interval.

S = Cells with less than 5 counts are suppressed as per the confidentiality policies of the Clinical Practice Research Datalink.

¥ Cytomegalovirus related infections are not included in the analysis due to a very low number of exposed events.
* Non-mutually exclusive categories.
** Matched on age, sex, date of cohort entry and duration of follow-up.
*** Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.

Table of Contents

Supplementary Figure 1. Restricted cubic spline of time since first infection on the risk of any dementia



The solid line shows the odds ratio and the dashed lines show the lower and upper bound of the 95% confidence interval. The curve begins at 2 years given the use of a 2-year lag period in the definition of exposure.

Supplementary Table 1. Read Codes for dementia

Read Code	Description
E00..12	Senile/presenile dementia
F110.00	Alzheimer's disease
Eu01.00	[X] Vascular dementia
Eu02z00	[X] Unspecified dementia
E00..11	Senile dementia
Eu00.00	[X] Dementia in Alzheimer's disease
E000.00	Uncomplicated senile dementia
Eu02z14	[X] Senile dementia NOS
Eu00z11	[X] Alzheimer's dementia unspecified
E004.11	Multi infarct dementia
Eu02500	[X] Lewy body dementia
E004.00	Arteriosclerotic dementia
Eu00200	[X] Dementia in Alzheimer's dis, atypical or mixed type
Eu02300	[X] Dementia in Parkinson's disease
E001.00	Presenile dementia
E041.00	Dementia in conditions EC
Eu00112	[X] Senile dementia, Alzheimer's type
Eu01100	[X] Multi-infarct dementia
Eu00z00	[X] Dementia in Alzheimer's disease, unspecified
Eu01300	[X] Mixed cortical and subcortical vascular dementia
Eu01z00	[X] Vascular dementia, unspecified
Eu01.11	[X] Arteriosclerotic dementia
F110100	Alzheimer's disease with late onset
Eu00100	[X] Dementia in Alzheimer's disease with late onset
Eu02.00	[X] Dementia in other diseases classified elsewhere
E012.11	Alcoholic dementia NOS
F110000	Alzheimer's disease with early onset
E002100	Senile dementia with depression
Eu10711	[X] Alcoholic dementia NOS
E002000	Senile dementia with paranoia
E003.00	Senile dementia with delirium
Eu00000	[X] Dementia in Alzheimer's disease with early onset
E004z00	Arteriosclerotic dementia NOS
Eu01200	[X] Subcortical vascular dementia
E002.00	Senile dementia with depressive or paranoid features
Eu01y00	[X] Other vascular dementia
E001z00	Presenile dementia NOS
Eu02000	[X] Dementia in Pick's disease
Eu02200	[X] Dementia in Huntington's disease

Read Code	Description
Eu02z16	[X] Senile dementia, depressed or paranoid type
E001200	Presenile dementia with paranoia
E012.00	Other alcoholic dementia
E001300	Presenile dementia with depression
E004000	Uncomplicated arteriosclerotic dementia
Eu02z13	[X] Primary degenerative dementia NOS
E001100	Presenile dementia with delirium
Eu02y00	[X] Dementia in other specified diseases classified elsewhere
E004300	Arteriosclerotic dementia with depression
Eu00113	[X] Primary degenerative dementia of Alzheimer's type, senile onset
Eu02z11	[X] Presenile dementia NOS
Eu01000	[X] Vascular dementia of acute onset
Eu00011	[X] Presenile dementia, Alzheimer's type
Eu00111	[X] Alzheimer's disease type 1
Eu02100	[X] Dementia in Creutzfeldt-Jakob disease
E004200	Arteriosclerotic dementia with paranoia
E001000	Uncomplicated presenile dementia
3AE4.00	GDS level 5 - moderately severe cognitive decline
E002z00	Senile dementia with depressive or paranoid features NOS
Eu01111	[X] Predominantly cortical dementia
E02y100	Drug-induced dementia
E004100	Arteriosclerotic dementia with delirium
Eu02400	[X] Dementia in human immunodeficiency virus [HIV] disease
Eu00012	[X] Primary degenerative dementia, Alzheimer's type, presenile onset
3AE5.00	GDS level 6 - severe cognitive decline
Fyu3000	[X] Other Alzheimer's disease
Eu00013	[X] Alzheimer's disease type 2
3AE6.00	GDS level 7 - very severe cognitive decline

Abbreviations: NOS, not otherwise specified; EC, elsewhere classified; GDS, Global Deterioration Scale.

Supplementary Table 2. Distribution of cases of Alzheimer's disease

Criterion	N (%*)
AD diagnosis with ≥1 prescription of an AD medication	7,563 (19)
Unspecified dementia diagnosis followed by ≥2 prescriptions of AD medications	4,372 (11)
≥2 records of AD diagnosis	2,434 (6)
AD diagnosis after dementia test or specialist referral or neuroimaging assessment	18,345 (45)
AD diagnosis with any dementia symptoms in any sequence	22,330 (55)

Abbreviations: AD, Alzheimer's disease.

* Criteria were not mutually exclusive.

Supplementary Table 3. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of Alzheimer's disease (stratified by age and sex)

Exposure	Cases, n (%)	Controls, n (%)	Crude OR (95% CI)	Adjusted* OR (95% CI)
<65 years old	**(n = 12,401)**	**(n = 495,129)**		
No infections	9,416 (75.9)	390,545 (78.9)	Reference	Reference
≥1 infection	2,985 (24.1)	104,584 (21.1)	1.20 (1.15 to 1.26)	1.05 (1.01 to 1.10)
≥65 years old	**(n = 28,054)**	**(n = 1,115,373)**		
No infections	22,989 (81.9)	930,929 (83.1)	Reference	Reference
≥1 infection	5,065 (18.1)	184,444 (16.9)	1.12 (1.08 to 1.15)	1.04 (1.01 to 1.08)
Male	**(n = 14,454)**	**(n = 573,799)**		
No infections	12,492 (86.4)	501,469 (87.3)	Reference	Reference
≥1 infection	1,962 (13.6)	72,330 (12.7)	1.08 (1.04 to 1.14)	0.99 (0.94 to 1.04)
Female	**(n = 26,001)**	**(n = 1,036,703)**		
No infections	19,913 (76.6)	820,005 (79.1)	Reference	Reference
≥1 infection	6,088 (23.4)	216,698 (20.9)	1.17 (1.13 to 1.20)	1.08 (1.04 to 1.11)

Abbreviations: OR, odds ratio; CI, confidence interval.

*Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.

Supplementary Table 4. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of Alzheimer's disease (sensitivity analyses)

Exposure	Cases n (%)	Controls n (%)	Crude OR (95% CI)	Adjusted*** OR (95% CI)
3-year lag*	**(n = 38,690)**	**(n = 1,540,372)**		
No infections	31,504 (81.4)	1,282,588 (83.2)	Reference	Reference
≥1 infection	7,186 (18.6)	257,784 (16.8)	1.14 (1.11 to 1.18)	1.05 (1.02 to 1.08)
5-year lag*	**(n = 34,717)**	**(n = 1,381,762)**		
No infections	29,093 (83.8)	1,179,010 (85.3)	Reference	Reference
≥1 infection	5,624 (16.2)	202,752 (14.7)	1.13 (1.10 to 1.17)	1.04 (1.01 to 1.07)
10-year lag*	**(n = 22,957)**	**(n = 912,548)**		
No infections	20,200 (88.0)	815,887 (89.3)	Reference	Reference
≥1 infection	2,757 (12.0)	96,661 (10.7)	1.16 (1.11 to 1.21)	1.08 (1.03 to 1.12)
Censoring on other dementia	**(n = 38,697)**	**(n = 1,521,962)**		
No infections	31,024 (80.2)	1,250,672 (82.1)	Reference	Reference
≥1 infection	7,673 (19.8)	271,290 (17.9)	1.15 (1.12 to 1.18)	1.05 (1.03 to 1.08)
Stricter exposure definition**	**(n = 40,455)**	**(n = 1,610,502)**		
No infections	32,725 (80.9)	1,334,297 (82.8)	Reference	Reference
≥1 infection	7,730 (19.1)	276,205 (17.2)	1.15 (1.12 to 1.18)	1.05 (1.02 to 1.08)

Abbreviations: OR, odds ratio; CI, confidence interval.

* Cases and controls with a duration of follow-up shorter than the duration of the lag were excluded from these analyses.

** Medical codes for pneumonia were restricted to those with a clear link to *Chlamydophila* (medical codes for other infections remained unchanged).

*** Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer,

oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.

Supplementary Table 5. Baseline characteristics of dementia cases and their matched controls[*]

Characteristic	Cases (n = 111,366)	Controls[a] (n = 4,419,042)
Age in years, mean (standard deviation)	81.7 (7.8)	81.7 (7.8)
Follow-up in years, mean (standard deviation)	10.2 (5.5)	10.2 (5.5)
Male sex	40,889 (36.7)	1,618,000 (36.7)
Body mass index in kg/m^2		
< 25	32,906 (26.7)	1,173,972 (24.2)
25-29	28,296 (23.0)	1,145,249 (23.7)
≥ 30	13,049 (10.6)	505,272 (10.6)
Unknown	49,021 (39.8)	1,995,597 (41.5)
Smoking		
Ever	38,805 (31.5)	1,353,577 (28.4)
Never	51,279 (41.6)	2,046,023 (42.0)
Unknown	33,188 (26.9)	1,420,490 (29.6)
Alcohol-related disorders	2,272 (2.0)	57,637 (1.3)
Arterial hypertension	37,124 (33.3)	1,420,360 (32.1)
Atrial fibrillation	4,234 (3.8)	139,220 (3.2)
Congestive heart failure	3,329 (3.0)	109,021 (2.5)
Coronary artery disease	18,668 (16.8)	652,770 (14.8)
Stroke or transient ischemic attack	6,782 (6.1)	190,246 (4.3)
Peripheral vascular disease	3,226 (2.9)	101,136 (2.3)
Dyslipidemia	11,197 (10.1)	395,995 (8.9)
Diabetes mellitus	9,924 (8.9)	307,333 (7.0)
Chronic kidney disease	2,762 (2.5)	96,164 (2.2)
Liver disease	397 (0.4)	13,510 (0.3)
Depression	14,686 (13.2)	447,041 (10.1)
Epilepsy	2,075 (1.9)	52,155 (1.2)
Parkinson's disease	1,037 (0.9)	17,187 (0.4)
Previous traumatic brain injury	7 (0.01)	92 (0.0)
Osteoporosis	3,937 (3.5)	132,425 (3.0)
Hypothyroidism	9,599 (8.6)	360,553 (8.1)
Cancer	9,989 (9.0)	388,671 (8.8)
Medications[b]		
Oral anticoagulants	10,908 (9.8)	384,841 (8.7)
Antiplatelet agents	50,534 (45.4)	1,588,400 (35.9)
Opioids	46,278 (41.6)	1,734,172 (39.2)
Lipid-lowering drugs	45,509 (40.9)	1,612,827 (36.4)
Antihypertensives[c]	75,157 (67.5)	2,941,945 (66.5)
Antipsychotics	17,831 (16.0)	349,644 (7.9)
Non-steroidal anti-inflammatory drugs	20,006 (18.0)	915,226 (20.6)
Antidepressants	37,848 (34.0)	801,189 (18.1)

[*] Numbers are presented as n (%) unless otherise specified.
[a] For controls, means and percentages were weighted by the inverse number of controls matched to each case.
[b] Measured in the two years prior to index date as a surrogate measure of overall health.
[c] Includes beta-blockers, thiazide diuretics, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, and calcium channel blockers

Supplementary Table 6. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of any dementia (overall and stratified by cumulative number of infections and time since first infection)

Exposure	Cases (n = 111,366) n (%)	Controls (n = 4,419,042) n (%)	Crude OR (95% CI)	Adjusted** OR (95% CI)
No infections	89,775 (80.6)	3,694,008 (83.6)	Reference	Reference
≥1 infection	21,591 (19.4)	725,034 (16.4)	1.25 (1.23 to 1.27)	1.08 (1.07 to 1.10)
Cumulative number of infections ('dose-response')				
No infections	89,775 (80.6)	3,694,008 (83.6)	Reference	Reference
1	13,813 (12.4)	470,825 (10.7)	1.22 (1.20 to 1.25)	1.09 (1.07 to 1.11)
2-3	5,693 (5.1)	186,323 (4.2)	1.28 (1.25 to 1.32)	1.08 (1.05 to 1.11)
>3	2,085 (1.9)	67,886 (1.5)	1.30 (1.25 to 1.36)	1.04 (0.99 to 1.09)
Time since first infection in years ('time-response')*				
No infections	89,775 (80.6)	3,694,008 (83.6)	Reference	Reference
2-3.9	5,336 (4.8)	169,413 (3.8)	1.30 (1.26 to 1.34)	1.13 (1.10 to 1.16)
4-6.9	5,839 (5.2)	196,758 (4.4)	1.24 (1.20 to 1.27)	1.08 (1.05 to 1.11)
7-10.9	5,211 (4.7)	181,880 (4.1)	1.20 (1.17 to 1.24)	1.05 (1.02 to 1.08)
11-30	5,205 (4.7)	176,983 (4.1)	1.24 (1.21 to 1.28)	1.08 (1.04 to 1.11)

Abbreviations: OR, odds ratio; CI, confidence interval.

* Given the use of a 2-year lag period in the definition of exposure, the minimum time since first infection was 2 years.

** Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors,

angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants. P values for trend were 0.11 for the dose-response analysis and <0.0001 for the time-response analysis.

Supplementary Table 7. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of any dementia (stratified by specific type of infection)

Exposure	Cases (n = 111,366) n (%)	Controls (n = 4,419,042) n (%)	Crude OR (95% CI)	Adjusted* OR (95% CI)
Specific type				
No infections	89,775 (80.6)	3,694,008 (83.6)	Reference	Reference
Urinary tract infections	15,950 (14.3)	534,616 (12.1)	1.23 (1.21 to 1.25)	1.08 (1.06 to 1.10)
Herpes	177 (0.2)	6,595 (0.2)	1.06 (0.92 to 1.23)	0.97 (0.83 to 1.13)
CMV related infections	S	65 (0.001)	0.62 (0.09 to 4.44)	0.52 (0.07 to 3.73)
Lyme disease	14 (0.01)	465 (0.01)	1.20 (0.71 to 2.05)	1.26 (0.74 to 2.15)
Gingivitis	585 (0.5)	18,199 (0.4)	1.28 (1.17 to 1.39)	1.11 (1.02 to 1.20)
Gastritis	4,154 (3.7)	134,536 (3.0)	1.24 (1.20 to 1.28)	1.06 (1.02 to 1.09)
Pneumonia	1,588 (1.4)	48,923 (1.1)	1.28 (1.22 to 1.35)	1.11 (1.05 to 1.16)
Candidiasis	2,111 (1.9)	69,992 (1.6)	1.20 (1.15 to 1.26)	1.01 (0.97 to 1.06)

Abbreviations: OR, odds ratio; CI, confidence interval; CMV, cytomegalovirus.

S = Cells with less than 5 counts are suppressed as per the confidentiality policies of the Clinical Practice Research Datalink.
* Non-mutually exclusive categories.

*Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.

Supplementary Table 8. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of any dementia (stratified by age and sex)

Exposure	Cases, n (%)	Controls, n (%)	Crude OR (95% CI)	Adjusted* OR (95% CI)
<65 years old	**(n = 25,927)**	**(n = 1,038,283)**		
No infections	19,591 (75.6)	829,292 (79.9)	Reference	Reference
≥1 infection	6,336 (24.4)	208,991 (20.1)	1.32 (1.28 to 1.36)	1.08 (1.04 to 1.11)
≥65 years old	**(n = 85,439)**	**(n = 3,380,759)**		
No infections	70,184 (82.2)	2,864,716 (84.5)	Reference	Reference
≥1 infection	15,255 (17.9)	516,043 (15.5)	1.22 (1.20 to 1.24)	1.08 (1.06 to 1.10)
Male	**(n = 40,889)**	**(n = 1,618,000)**		
No infections	35,125 (85.9)	1,427,426 (88.2)	Reference	Reference
≥1 infection	5,764 (14.1)	190,574 (11.8)	1.24 (1.20 to 1.27)	1.07 (1.04 to 1.10)
Female	**(n = 70,477)**	**(n = 2,801,042)**		
No infections	54,650 (77.5)	2,266,582 (80.9)	Reference	Reference
≥1 infection	15,827 (22.5)	534,460 (19.1)	1.25 (1.23 to 1.27)	1.09 (1.08 to 1.12)

Abbreviations: OR, odds ratio; CI, confidence interval.

*Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.

Supplementary Table 9. Crude and adjusted odds ratios for the association between infectious disease burden and the risk of any dementia (sensitivity analyses)

Exposure	Cases n (%)	Controls n (%)	Crude OR (95% CI)	Adjusted*** OR (95% CI)
3-year lag*	**(n = 103,469)**	**(n = 4,105,728)**		
No infections	84,732 (81.9)	3,470,253 (84.5)	Reference	Reference
≥1 infection	18,737(18.1)	635,475 (15.5)	1.23 (1.20 to 1.25)	1.07 (1.05 to 1.09)
5-year lag*	**(n = 88,757)**	**(n = 3,521,396)**		
No infections	74,709 (84.2)	3,038,510 (86.2)	Reference	Reference
≥1 infection	14,048 (15.8)	482,886 (13.8)	1.20 (1.17 to 1.22)	1.05 (1.03 to 1.07)
10-year lag*	**(n = 52,989)**	**(n = 2,098,289)**		
No infections	46,732 (88.2)	1,884,398 (89.7)	Reference	Reference
≥1 infection	6,257 (11.8)	213,891 (10.3)	1.19 (1.16 to 1.22)	1.06 (1.03 to 1.09)
Stricter exposure definition**	**(n = 111,366)**	**(n = 4,419,042)**		
No infections	90,800 (81.5)	3,727,314 (84.3)	Reference	Reference
≥1 infection	20,566 (18.5)	691,728 (15.7)	1.24 (1.22 to 1.26)	1.08 (1.06 to 1.10)

Abbreviations: OR, odds ratio; CI, confidence interval.

* Cases and controls with a duration of follow-up shorter than the duration of the lag were excluded from these analyses.

** Medical codes for pneumonia were restricted to those with a clear link to *Chlamydophila* (medical codes for other infections remained unchanged).

*** Adjusted for body mass index, smoking, alcohol-related disorders, arterial hypertension, atrial fibrillation, congestive heart failure, coronary artery disease, stroke or transient ischemic attack, peripheral vascular disease, dyslipidemia, diabetes mellitus, chronic kidney disease, liver disease, depression, epilepsy, Parkinson's disease, traumatic brain injury, osteoporosis, hypothyroidism, cancer, oral anticoagulants, antiplatelet agents, opioids, lipid-lowering drugs, beta-blockers, thiazides, angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium channel blockers, antipsychotics, non-steroidal anti-inflammatory drugs, and antidepressants.