Supplementary materials

Syntax used for database searches

Medline (Pubmed Interface):

("dementia"[MeSH Terms] OR "dementia"[All Fields] OR "Alzheimer disease"[MeSH Terms] OR "Alzheim*"[All Fields]) AND (("time"[MeSH Terms] OR "time"[All Fields]) OR ("trends"[All Fields] OR "trend*"[All Fields]) OR "secular"[All Fields] OR "change*"[All Fields]) AND ("epidemiology"[MeSH Terms] OR "epidemiology"[All Fields] OR "incidence"[MeSH Terms] OR "incidence"[All Fields] OR "cohort"[All Fields]) AND ("1990/01/01"[PDAT] : "2017/06/06"[PDAT]) AND ("humans"[MeSH Terms] AND (English[lang] OR German[lang]) AND (medline[sb] OR jsubsetaim[text] OR jsubsetn[text]) AND ("middle aged"[MeSH Terms] OR "aged"[MeSH Terms]))

Web of Science (Web of Science interface):

TOPIC: (((dementia OR Alzheimer disease OR Alzheim*) AND (time OR trends OR trend* OR secular OR change*) AND (epidemiology OR incidence OR cohort)))
Refined by: LANGUAGES: (ENGLISH OR GERMAN)
Timespan: 1990-2017. **Indexes:** SCI-EXPANDED, SSCI, A&HCI, ESCI.

Carried out June 6th, 2017.

References of excludes studies after full-text analysis

Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Time trends of incidence of age-associated diseases in the US elderly population: medicare-based analysis. *Age Ageing*. 2013;42(4):494-500. doi:10.1093/ageing/aft032.

Bronskill SE, Ng R, Yates E, et al. Trends in dementia prevalence, incidence and health system costs: a population-based analysis from Ontario, Canada. *Alzheimer's & Dementia*. 2016;12(7):P815-P816. doi:10.1016/j.jalz.2016.06.1651.

Cerasuolo JO, Cipriano LE, Sposato LA, et al. Population-based stroke and dementia incidence trends: Age and sex variations. *Alzheimers Dement*. 2017. doi:10.1016/j.jalz.2017.02.010.

Doblhammer G, Fink A, Zylla S, Willekens F. Compression or expansion of dementia in Germany? An observational study of short-term trends in incidence and death rates of dementia between 2006/07 and 2009/10 based on German health insurance data. *Alz Res Ther*. 2015;7(1):66. doi:10.1186/s13195-015-0146-x.

Fujishima M, Kiyohara Y. Incidence and risk factors of dementia in a defined elderly Japanese population - The Hisayama Study. *Alzheimer's Disease: Vascular Etiology and Pathology*. 2002;977:1-8.

Ganguli M, Dodge HH, Chen P, Belle S, DeKosky ST. Ten-year incidence of dementia in a rural elderly US community population: the MoVIES Project. *Neurology*. 2000;54(5):1109-1116.

Hagnell O, Franck A, Grasbeck A, et al. Senile dementia of the Alzheimer type in the Lundby Study 1. A prospective, epidemiologic study of incidence and risk during the 15 years 1957-1972. *Eur Arch Psychiatry Clin Neurosci*. 1991;241(3):159-164. doi:10.1007/BF02219715.

Hagnell O, Franck A, Grasbeck A, et al. Vascular dementia in the Lundby Study 1. A prospective, epidemiologic study of incidence and risk during the 15 years 1957-1972. *Neuropsychobiology*. 1992;26(1-2):43-49. doi:10.1159/000118895.

Hung Y-N, Kadziola Z, Brnabic AJM, et al. The epidemiology and burden of Alzheimer's disease in Taiwan utilizing data from the National Health Insurance Research Database. *Clinicoeconomics and Outcomes Research*. 2016;8:387-395. doi:10.2147/CEOR.S93323.

Jorgensen TSH, Torp-Pedersen C, Gislason GH, Andersson C, Holm E. Time trend in Alzheimer diagnoses and the association between distance to an Alzheimer clinic and Alzheimer diagnosis. *Eur J Public Health*. 2015;25(3):522-527. doi:10.1093/eurpub/cku118.

Kokmen E, Beard CM, Petersen RC, Ballard DJ, Kurland LT. Incidence, prevalence and risk factor-analysis of Alzheimers-Disease in Rochester, Minnesota – Long-term trends over 20 years. *Neurobiol Aging*. 1990;11(3):293.

Kokmen E, Beard CM, OBrien PC, Offord KP, Kurland LT. Is the incidence of dementing illness changing – a 25-year time trend study in Rochester, Minnesota (1960-1984). *Neurology*. 1993;43(10):1887-1892.

Kokmen E, Beard CM, O'Brien PC, Kurland LT. Epidemiology of dementia in Rochester, Minnesota. *Mayo Clin Proc*. 1996;71(3):275-282. doi:10.1016/S0025-6196(11)63967-0.

Kosteniuk JG, Morgan DG, O'Connell ME, et al. Incidence and prevalence of dementia in linked administrative health data in Saskatchewan, Canada: a retrospective cohort study. *BMC Geriatr*. 2015;15:73. doi:10.1186/s12877-015-0075-3.

Kosteniuk JG, Morgan DG, O'Connell ME, et al. Simultaneous temporal trends in

dementia incidence and prevalence, 2005-2013: a population-based retrospective cohort study in Saskatchewan, Canada. *Int Psychogeriatr*. 2016;28(10):1643-1658. doi:10.1017/S1041610216000818.

Qiu C, Strauss E von, Backman L, Winblad B, Fratiglioni L. Twenty-year changes in dementia occurrence suggest decreasing incidence in central Stockholm, Sweden. *Neurology*. 2013;80(20):1888-1894.

Rocca WA, Cha RH, Waring SC, Kokmen E. Incidence of dementia and Alzheimer's disease: a reanalysis of data from Rochester, Minnesota, 1975-1984. *Am J Epidemiol*. 1998;148(1):51-62.

Rocca WA, Kokmen E, Beard CM. Time trends in the incidence of dementia and Alzheimers-disease in Rochester, Minnesota – a reappraisal. *Ann Neurol*. 1994;36(2):269-270.

Sacuiu S, Gustafson D, Sjogren M, et al. Secular changes in cognitive predictors of dementia and mortality in 70-year-olds. *Neurology*. 2010;75(9):779-785. doi:10.1212/WNL.0b013e3181f0737c.

Sposato LA, Kapral MK, Fang J, et al. Declining Incidence of Stroke and Dementia: Coincidence or Prevention Opportunity? *JAMA Neurol*. 2015;72(12):1529-1531. doi:10.1001/jamaneurol.2015.2816.

van Bussell EF, Richard E, Arts DL, et al. Dementia incidence trend over 1992-2014 in the Netherlands: Analysis of primary care data. *PLoS Med*. 2017;14(3). doi:10.1371/journal.pmed.1002235.

Wancata J, Kaup B, Krautgartner M. Increasing number of people suffering from dementia between the years 1951 and 2050 in Austria. *Wien Klin Wochenschr*. 2001;113(5-6):172-180.

Wang H-W, Chen D-R, Yu H-W, Chen Y-M. Forecasting the Incidence of Dementia

and Dementia-Related Outpatient Visits With Google Trends: Evidence From Taiwan. *J Med Internet Res*. 2015;17(11). doi:10.2196/jmir.4516.

Whalley LJ. Spatial Distribution and Secular Trends in the Epidemiology of Alzheimer's Disease. *Neuroimaging Clinics of North America*. 2012;22(1):1-+. doi:10.1016/j.nic.2011.11.002.

Adapted risk of bias assessment tool by Hoy et al., 2012[1]

Risk of bias item	Criteria for answers	Additional notes and examples	Adaptions made
External Validity			
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	**Yes (LOW RISK):** The study's target population was a close representation of the national population. **No (HIGH RISK)**: The study's target population was clearly NOT representative of the national population.	The **target population** refers to the group of people or entities to which the results of the study will be generalised. Examples: • The study was a national health survey of people 15 years and over and the sample was drawn from a list that included all individuals in the population aged 15 years and over. The answer is: **Yes (LOW RISK).** • The study was conducted in one province only, and it is not clear if this was representative of the national population. The answer is: **No (HIGH RISK).** • The study was undertaken in one village only and it is clear this was not representative of the national population. The answer is: **No (HIGH RISK).**	N/A
2. Was the sampling frame a **true or close representation** of the target population?	**Yes (LOW RISK):** The sampling frame was a **true or close** representation of the target population. **No (HIGH RISK)**: The sampling frame was NOT a **true or close** representation of the target population.	The **sampling frame** is a list of the sampling units in the target population and the study sample is drawn from this list. Examples: • The sampling frame was a list of almost every individual within the target population. The answer is: **Yes (LOW RISK)**. • The cluster sampling method was used and the sample of clusters/villages was drawn from a list of all villages in the target population. The answer is: **Yes (LOW RISK)**. • The sampling frame was a list of just one particular ethnic group within the overall target population, which comprised many groups. The answer is: **No (HIGH RISK).**	N/A
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	**Yes (LOW RISK):** A census was undertaken, OR, some form of random selection was used to select the sample (e.g. simple random sampling, stratified random sampling, cluster sampling, systematic sampling). **No (HIGH RISK)**: A census was NOT undertaken, AND some form of random selection was NOT used to select the sample.	A census collects information from every unit in the sampling frame. In a survey, only part of the sampling frame is sampled. In these instances, random selection of the sample helps minimise study bias. Examples: • The sample was selected using simple random sampling. The answer is: **Yes (LOW RISK)**. • The target population was the village and every person in the village was sampled. The answer is: **Yes (LOW RISK)**. • The nearest villages to the capital city were selected in order to save on the cost of fuel. The answer is: **No (HIGH RISK).**	N/A
4. Was the likelihood of **non-response bias minimal?**	**Yes (LOW RISK):** The response rate for the study was >/=75%, OR, an analysis was performed that showed no significant difference in relevant demographic	Examples: • The response rate was 68%; however, the researchers did an analysis and found no significant difference between responders and non-responders in terms of age, sex, occupation and	N/A

	characteristics between responders and nonresponders **No (HIGH RISK)**: The response rate was <75%, and if any analysis comparing responders and non-responders was done, it showed a significant difference in relevant demographic characteristics between responders and non-responders	socioeconomic status. The answer is: **Yes (LOW RISK)**. • The response rate was 65% and the researchers did NOT carry out an analysis to compare relevant demographic characteristics between responders and non-responders. The answer is: **No (HIGH RISK)**. • The response rate was 69% and the researchers did an analysis and found a significant difference in age, sex and socio-economic status between responders and non-responders. The answer is: **No (HIGH RISK)**.	
Internal Validity			
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	**Yes (LOW RISK):** All data were collected directly from the subjects. **No (HIGH RISK)**: In some instances, data were collected from a proxy.	A proxy is a representative of the subject. Examples: • All eligible subjects in the household were interviewed separately. The answer is: **Yes (LOW RISK)**. • A representative of the household was interviewed and questioned about the presence of low back pain in each household member. The answer is: **No (HIGH RISK)**.	N/A
6. Was an acceptable case definition used in the study?	**Yes (LOW RISK):** An acceptable case definition was used. **No (HIGH RISK)**: An acceptable case definition was NOT used.	Examples: • For a study on low back pain, the following case definition was used: "Low back pain is defined as activity-limiting pain lasting more than one day in the area on the posterior aspect of the body • from the bottom of the 12th rib to the lower gluteal folds." The answer is: **Yes (LOW RISK)**. • For a study on back pain, there was no description of the specific anatomical location „back" referred to. The answer is: **No (HIGH RISK)**. • For a study on osteoarthritis, the following case definition was used: "Symptomatic osteoarthritis of the hip or knee, radiologically confirmed as Kellgren-Lawrence grade 2-4". The answer is: **Yes (LOW RISK).**	N/A
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	**Yes (LOW RISK):** The study instrument had been shown to have reliability and validity (if this was necessary), e.g. test-retest, piloting, validation in a previous study, etc. **No (HIGH RISK)**: The study instrument had NOT been shown to have reliability or validity (if this was necessary).	Examples: • The authors used the COPCORD questionnaire, which had previously been validated. They also tested the inter-rater reliability of the questionnaire. The answer is: **Yes (LOW RISK)**. • The authors developed their own questionnaire and did not test this for validity or reliability. The answer is: **No (HIGH RISK)**.	N/A
8. Was the **same mode of data collection** used for all subjects?	**Yes (LOW RISK):** The same mode of data collection was used for all subjects. **No (HIGH RISK)**: The same mode of data collection was NOT used	The mode of data collection is the method used for collecting information from the subjects. The most common modes are face-to-face interviews, telephone interviews and self-administered questionnaires. Examples:	N/A

	for all subjects.	• All eligible subjects had a face-to-face interview. The answer is: **Yes (LOW RISK)**. • Some subjects were interviewed over the telephone and some filled in postal questionnaires. The answer is: **No (HIGH RISK)**.	
9. Was the **length of the shortest prevalence period** for the parameter of interest appropriate?	**Yes (LOW RISK):** The shortest prevalence period for the parameter of interest was appropriate (e.g. point prevalence, one-week prevalence, one-year prevalence). **No (HIGH RISK)**: The shortest prevalence period for the parameter of interest was not appropriate (e.g. lifetime prevalence).	The prevalence period is the period that the subject is asked about e.g. "Have you experienced low back pain over the previous year?" In this example, the prevalence period is one year. The longer the prevalence period, the greater the likelihood of the subject forgetting if they experienced the symptom of interest (e.g. low back pain). Examples: • Subjects were asked about pain over the past week. The answer is: **Yes (LOW RISK)**. • Subjects were only asked about pain over the past three years. The answer is: **No (HIGH RISK)**.	9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate? **Yes (LOW RISK):** The paper reported length of follow-up interval, length of study time, and time between cohorts, so that it was comprehensible for which time interval incidence rates, cumulative incidence or incidence rate change were established. Follow-up periods in studies were appropriate so that new cases would be relibaley detectable, and not biased through too long intervals that might lead to a loss of cases due to death, attrition, etc. **No (HIGH RISK):** The study did not properly report time intervals for which incidence measures apply. Follow-up periods were too short (e.g., less than three months that might not be enough time to reliably detect new cases as disease evolution might take longer), or too long (e.g., follow-up after >10 years, leading to an underestimation of new cases due to competing risk of death, loss-to-follow-up, etc.).
10. Were the **numerator(s) and denominator(s)** for the parameter of interest appropriate?	**Yes (LOW RISK):** The paper presented appropriate numerator(s) AND denominator(s) for the parameter of interest (e.g. the prevalence of low back pain). **No (HIGH RISK)**: The paper did present numerator(s) AND denominator(s) for the parameter of interest but one or more of these were inappropriate.	There may be errors in the calculation and/or reporting of the numerator and/or denominator. Examples: • There were no errors in the reporting of the numerator(s) AND denominator(s) for the prevalence of low back pain. The answer is: **Yes (LOW RISK)**. • In reporting the overall prevalence of low back pain (in both men and women), the authors accidentally used the population of women as the denominator rather than the combined population. The answer is: **No (HIGH RISK).**	10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate change**? **Yes (LOW RISK):** The paper presented an appropriate, i.e. lege artis methodological approach to calculate cumulative incidence or incidence rate (e.g., person-year method, Poisson, annual incidence rate, etc.) and incidence rate change (incidence rate ratio, hazard ratio, odds ratio, etc.). **No (HIGH RISK):** The paper did not describe or use appropriate methods to establish incidence rate or incidence rate change, so that it was not clear how measures were calculated.
11. Summary item on the overall risk of study bias			
LOW RISK OF BIAS: Further research is very unlikely to change our confidence in the estimate. **MODERATE RISK OF BIAS:** Further research is likely to have an important impact on our confidence in the estimate and may change the estimate. **HIGH RISK OF BIAS:** Further research is very likely to have an important impact on our confidence in the estimate and is likely to change the estimate.			
Abbreviations: N/A = not applicable			

Item-based risk of bias assessment and justification for judgment in individual studies.

Study 1

Author(s): Walter A. Rocca , Ronald C. Petersen, David S. Knopman, Liesi E. Hebert, Denis A. Evans, Kathleen S. Hall, Sujuan Gao, Frederick W. Unverzagt, Kenneth M. Langa, Eric B. Larson, Lon R. White

Year of publication: 2011

Name of paper/study: Chicago Health and Aging Project (CHAP)

Reference: Rocca WA, Petersen RC, Knopman DS, et al. Trends in the incidence and prevalence of Alzheimer's disease, dementia, and cognitive impairment in the United States. *Alzheimers Dement*. 2011;7(1):80–93.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the *older* national population in relation to relevant variables, e.g. age, sex, occupation?	No (= high risk)	The sample was population-based and bi-racial. However, the population was only one community. It is not clear if this sample was representative of the national population.
2. Was the sampling frame a **true or close representation** of the target population?	Yes (= low risk)	The target population was every person in the community – all of them were sampled.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	The sampling frame comprised all residents aged 65 years in a defined geographic area.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	Yes	The response rate at baseline was 78.7%.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	The analyses used data from subjects receiving clinical evaluation.
6. Was an **acceptable case definition** used in the study?	Yes	Diagnosis of dementia required loss of cognitive function by the neurologist's assessment and impairment in two or more functions on cognitive performance tests. The diagnosis of Alzheimer's disease was by criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) for probable Alzheimer's disease.
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	Cognitive measures included validated neuropsychological tests (Verbal Fluency Test, Boston Naming Test, Mini-Mental State Examination, Word List Memory, Word List Recall and Word List Recognition from the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) battery, Logical Memory and Digit Span subtests of the Wechsler Memory Scale

		– Revised, East Boston Memory Test and modified versions of the Symbol Digit Modalities Test, Judgment of Line Orientation, Complex Ideational Material, Number Comparison, Digit Ordering, Standard Progressive Matrices, and National Adult Reading Test.
8. Was the **same mode of data collection** used for all subjects?	Yes	Each cycle consists of two phases: (1) an in-home interview, requiring one to one-and-a-half hours to complete, of all participants; and (2) clinical evaluation for Alzheimer's disease, other dementing conditions, and other medical conditions, of a stratified random sample of persons. Moreover, clinical evaluations for incidence of Alzheimer's disease were uniform throughout the study.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	Data were collected in 3-year cycles. There have been a total of four cohort intervals. Evaluations for incident disease have been performed approximately continuously since the beginning of the second cycle in 1997, for a total of 11 years.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Unclear	Cumulative incidence and person-years were not reported; only number of new cases during study period.
Summary item on the overall risk of study bias		
LOW RISK OF BIAS: Further research is very unlikely to change our confidence in the estimate.		

Study 2

Author(s): E.M.C. Schrijvers, B.F.J. Verhaaren, P.J. Koudstaal, A. Hofman, M.A. Ikram, M.M.B. Breteler

Year of publication: 2012

Name of paper/study: The Rotterdam Study

Reference: Schrijvers EM, Verhaaren BF, Koudstaal PJ, Hofman A, Ikram MA, Breteler MM. Is dementia incidence declining?: trends in dementia incidence since 1990 in the Rotterdam Study. *Neurology*. 2012;78(19):1456–1463.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the *older* national population in relation to relevant variables, e.g. age, sex, occupation?	No	The study was conducted in one district only, and it is not clear if this was representative of the national population.
2. Was the sampling frame a **true or close representation** of the target population?	Yes	The sampling comprised all inhabitants aged 55 years and older of Ommoord, a district of Rotterdam.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	The target population was every person in the community – all of them were sampled.

4. Was the likelihood of **non-response bias minimal (≥75%)?**	No	The overall response rate for all three cycles at baseline was 72%. For the main cohorts compared, it was 78,15% in the 1990 cohort, and 67,33% in the 2000 cohort. There was no report of analysis to compare responders to non-responders.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	Participants were interviewed at home (2h) and then had an extensive set of examinations (a total of 5 hours) in a specially built research facility in the centre of their district.
6. Was an **acceptable case definition** used in the study?	Yes	The diagnosis of dementia was made in accordance with internationally accepted criteria (*DSM-III-R*) by a panel of a neurologist, neuropsychologist, and research physician.
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	Two brief tests of cognition (Mini-Mental State Examination [MMSE] and Geriatric Mental State schedule [GMS] organic level) were used to screen all subjects. Screen-positives (MMSE score <26 or GMS organic level >0) underwent the Cambridge examination for mental disorders of the elderly.
8. Was the **same mode of data collection** used for all subjects?	Yes	Important strengths of the study are […] the equal assessment of dementia and vascular risk factors in both subcohorts.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	Participants contributed person-years for a maximum of 5 years after baseline. The length of time interval between cohorts was 10 years.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Yes	Incidence rates per 1,000 person-years and incidence rate ratios (IRR) were calculated using Poisson regression models for the 2 subcohorts in total, in 10-year age strata, and for men and women separately. All analyses were adjusted for age, and an additional adjustment was made for age squared, to make sure the effect of age was adequately adjusted for.
Summary item on the overall risk of study bias		
LOW RISK OF BIAS: Further research is very unlikely to change our confidence in the estimate.		

Study 3

Author(s): Sujuan Gao, Adesola Ogunniyi, Kathleen S. Hall, Olusegun Baiyewu, Frederick W. Unverzagt, Kathleen A. Lane, Jill R. Murrell, Oye Gureje, Ann M. Hake, Hugh C. Hendrie

Year of publication: 2016

Name of paper/study: Indianapolis-Ibadan Dementia Project (IIDP)

Reference: Gao S, Ogunniyi A, Hall KS, et al. Dementia incidence declined in African-Americans but not in Yoruba. *Alzheimers Dement*. 2016;12(3):244–251.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	No	Community-based cohort of a specific population (African-Americans). Thus, it is not known whether the rates reported are generalizable to other population, let alone the older national population.
2. Was the sampling frame a **true or close representation** of the target population?	No	The sampling frame was a list of just one particular ethnic group within the overall target population.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	In 1992, interviewers went door-to-door to randomly sampled addresses to invite African-Americans (self-identified) aged ≥65 years to participate. In 2001, subjects were randomly selected from Medicare records.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	No	Adequate response rate in the first cohort in 1992 (85.7%), but lower response rate in the second cohort in 2001 (44.0%). Differences between responders and non-responders reported. Therefore, selection bias based on differential refusal rates cannot be entirely ruled out.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	A two-stage design was used at each evaluation with in-home cognitive and functional evaluations for all participants followed by a full diagnostic workup of selected participants based on the performance of stage 1 cognitive tests.
6. Was an **acceptable case definition** used in the study?	Yes	Dementia was diagnosed with both the Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition and International Classification of Diseases, 10th Revision criteria. AD was diagnosed using criteria proposed by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS)/ Alzheimer's Disease and Related Disorder Association (ADRDA).
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	The community screening interview for dementia (CSID) was developed by the study group. However, the validity and reliability (inter-rater reliability) have

		been tested.
8. Was the **same mode of data collection** used for all subjects?	No	Not all participants received evaluation at the second stage. Therefore, dementia incidence was estimated in those individuals.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	A baseline evaluation was followed by regular evaluations scheduled 2–3 years apart in both populations using identical assessment instruments. Participants in the 1992 cohorts were evaluated for up to seven times, in 1992, 1995, 1998, 2001, 2004, 2007, and 2009. Participants in the 2001 cohort were evaluated for up to four times, in 2001, 2004, 2007, and 2009.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Yes	The person-years method was used to estimate age-specific incident rates from nondementia to dementia. ForIncident dementia rates for a specific age group were derived as the total estimated number of incident dementia cases divided by the total estimated person-years at risk for the age group. For estimated age-specific incidence rates, 95% confidence intervals (CIs) were derived using the inverse of the c2-distributions evaluated with the predicted dementia cases to approximate Poisson distributions. Age-standardized overall incident rates were obtained by applying the estimated age-specific rates to the age distribution of African-American residents in Marion County (Indianapolis) observed in the 2000 census. Significant difference between two rates in comparison was established by nonoverlapping 95%CIs.
Summary item on the overall risk of study bias		
MODERATE RISK OF BIAS: Further research is likely to have an important impact on our confidence in the estimate and may change the estimate.		

Study 4

Author(s): Leslie Grasset, Carol Brayne, Pierre Joly, Hélène Jacqmin-Gadda, Karine Peres, Alexandra Foubert-Samier, Jean-François Dartigues, Catherine Helmer

Year of publication: 2016

Name of paper/study: Personnes Agées Quid (PAQUID) & Three-City (3C)

Reference: Grasset L, Brayne C, Joly P, et al. Trends in dementia incidence: evolution over a 10-year period in France. *Alzheimers Dement*. 2016;12(3):272–280.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	No	PAQUID: The study was conducted in 75 perishes of Southwestern France. It is not clear if this was representative of the national population. 3C: The study was undertaken in 3 French cities (Bordeaux, Dijon, Montpellier). It is clear that the sample is an imperfect representation of the population. Participants differ from the general population in age, sex, and socioeconomic level distributions.
2. Was the sampling frame a **true or close representation** of the target population?	Yes	PAQUID: A sample of individuals on the electoral rolls was randomly selected from all community residents aged 65 years and over living at home in 75 parishes of Gironde and Dordogne (southwestern France). 3C: Stratified sampling was used based electoral rolls.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	PAQUID: A sample of individuals on the electoral rolls was randomly selected from all community residents aged 65 years and over living at home in 75 parishes of Gironde and Dordogne (southwestern France). 3C: Stratified sampling was used based electoral rolls.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	No	PAQUID: The response was 60%. A random sample of refusals did not differ from participants for age, sex, and educational level. 3C: The response rate was 39%. The lower response rate of the 2000s cohort will potentially have led to selection bias, with participants differing somewhat from the population in age, sex, and socioeconomic level distribution.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	For both cohorts, a standardized questionnaire assessing socio-demographic, medical, cognitive, and functional data was administered by trained neuropsychologists during face-to-face interviews, at baseline, and at

		each follow-up.
6. Was an **acceptable case definition** used in the study?	Yes	Two definitions of dementia have been used: a clinical diagnosis and an algorithmic diagnosis, both considered as nonreversible diagnosis. The clinical diagnosis was made after a three-step procedure for both 1990s and 2000s populations. The first step was a cognitive evaluation made by the neuropsychologist through a series of psychometric tests. Participants who were suspected of dementia, based on their neuropsychological performances or decline relative to a previous examination, were then examined by a senior neurologist. The diagnosis of dementia was based on the Diagnostic and Statistical Manual of Mental Disorders - Third Edition, Revised and the Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition criteria. The algorithmic diagnosis was based on cognitive and functional assessments, using the mini mental state examination (MMSE) and the four instrumental activities of daily living (IADL; ability to use the telephone, transportation, responsibility for medications, and ability to manage its budget).
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	Besides clinical evaluation, a series of psychometric tests was applied, incl. the Mini-Mental State Examination, Trail Making Test, Delayed Recall, Digit Span Test.
8. Was the **same mode of data collection** used for all subjects?	Yes	Important strengths of the study are the comparison of two large independent cohorts, with identical design and procedures of data collection.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	Follow-up took place after 3, 5, 8, and 10 years for PAQUID and 2, 4, 7, and 10 years for 3C. For each of the two populations, data from the first 10 years of follow-up were analyzed (1988–1989 to 1998–1999 for the 1990s population and 1999–2001 to 2009–2010 for the 2000s population).
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Unclear	It is not fully traceable how incidence rates were established. Incidence rates are not presented numerically, but graphically only. Incidence analyses were performed using a multi-state illness-death model. HR between the two cohorts were then calculated, adjusted for age and sex.
Summary item on the overall risk of study bias		
MODERATE RISK OF BIAS: Further research is likely to have an important impact on our confidence in the estimate and may change the estimate.		

Study 5

Author(s): F.E. Matthews, B.C.M. Stephan, L. Robinson, C. Jagger, L.E. Barnes, A. Arthur, C. Brayne, Cognitive Function and Ageing Studies (CFAS) Collaboration

Year of publication: 2016

Name of paper/study: UK Medical Research Council Cognitive Function and Ageing Study (MCR CFAS I & II)

Reference: Matthews FE, Stephan BC, Robinson L, et al. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. *Nat Commun*. 2016;7:11398.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	Yes	Multi-centre approach with rural/ urban and north/south areas. Baseline interviews in complete populations samples aged 65 years and over. Three of the areas were selected on the basis of representing the range of prevalence estimates. Comprehensive analyses concerning non-response etc. was done.
2. Was the sampling frame a **true or close representation** of the target population?	Yes	Baseline interviews in complete population samples aged 65 years and over.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	Both studies used the UK system of primary care registration, which provides the most robust population sampling frame (including institutions) and allows true geographical sampling held steady over time.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	Yes	The response rate in CFAS II was 80%, in CFAS II 56%. Adjustment for initial non-response made little difference to the estimate of incidence.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	An introductory letter from the general practitioner was followed by a visit from a named study interviewer. At 2 year follow-up all those respondents who had provided an interview and were still alive were re-approached. Each individual was visited up to three times to maximize response at each time period.
6. Was an **acceptable case definition** used in the study?	Yes	Algorithmic approach to dementia diagnosis, AGECAT, Mini-Mental-State-Examination (MMSE), based on the Geriatric Mental State (GMS) examination.
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	The GMS-AGECAT has been validated against internationally accepted earlier diagnostic criteria (DSM-III-R).
8. Was the **same mode of data collection** used for all subjects?	Yes	CFAS I and CFAS II had identical designs, methods and diagnostic approach apart from the simplification

		of design from two stage to one stage at baseline and incidence phase through combination of screen and assessment interviews, though the CFAS I interview was completely included in CFAS II.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	CFAS I: baseline between 1990-93, CFAS II: baseline between 2008-2011; follow-up in each study after 2 years.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Yes	Incidence rates per 1,000 person years per age and sex in CFAS I and in CFAS II, respectively. The reduction in incidence was calculated as an incidence rate ratio (IRR) 0.8 (95% CI: 0.6–1.0, P¼0.08, Poisson regression, N=10,444).
Summary item on the overall risk of study bias		
LOW RISK OF BIAS: Further research is very unlikely to change our confidence in the estimate.		

Study 6

Author(s): Claudia L. Satizabal, Alexa S. Beiser, Vincent Chouraki, Geneviève Chêne, Carole Dufouil, Sudha Seshadri

Year of publication: 2016

Name of paper/study: Framingham Heart Study (FHS)

Reference: Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of dementia over three decades in the Framingham Heart Study. *N Engl J Med*. 2016;374(6):523–532.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	No	One of the limitations of the Framingham Heart Study is that the participants are overwhelmingly of European ancestry. Therefore, the findings would need to be replicated in groups that include a larger number of participants of other races and ethnic backgrounds.
2. Was the sampling frame a **true or close representation** of the target population?	No	The final cohort was healthier than the general population.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	No	It was planned to obtain the sample via random selection; however, volunteers had to be added to obtain the desired sample size.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	No	The initial response rate of the selected sample was 68,7%. This rate plus the inclusion of volunteers may have introduced biases beyond the obvious one of a healthier cohort than the general population.
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	No	While most participants were interviewed face-to-face,

		some participants who moved away may be brought to review based only on assessments by their treating physician. A few persons are brought to dementia review based only on their medical record and, if available, a family interview.
6. Was an **acceptable case definition** used in the study?	Yes	Diagnosis of dementia is based on DSM-IV criteria.
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	Neuropsychological test battery included reliable and valid tests, e.g. MMSE, Digit Span, Trail-making Test A and B etc.
8. Was the **same mode of data collection** used for all subjects?	No	While most participants were interviewed face-to-face, some participants who moved away may be brought to review based only on assessments by their treating physician. A few persons are brought to dementia review based only on their medical record and, if available, a family interview.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	Cumulative incidence of dementia (per 100 persons) over a period of 5 years in 4 non-overlapping epochs over three decades.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence rate** change?	Yes	Cox proportional-hazards models, adjusted for age at entry and sex, were used to compare the incidence of dementia across the four epochs. For each epoch, the 5-year cumulative hazard rates were calculated, which represent the cumulative incidence of dementia per 100 persons over a period of 5 years. Hazard ratios which represent the incidence of dementia during each epoch relative to the incidence during the first epoch are reported.
Summary item on the overall risk of study bias		
HIGH RISK OF BIAS: Further research is very likely to have an important impact on our confidence in the estimate and is likely to change the estimate.		

Study 7

Author(s): Tomoyuki Ohara, Jun Hata, Daigo Yoshida, Naoko Mukai, Masaharu Nagata, Toru Iwaki, Takanari Kitazono, Shigenobu Kanba, Yutaka Kiyohara, Toshiharu Ninomiya

Year of publication: 2017

Name of paper/study: The Hisayama Study

Reference: Ohara T, Hata J, Yoshida D, et al. Trends in dementia prevalence, incidence, and survival rate in a Japanese community. *Neurology*. 2017;88(20):1925–1932.

Risk of bias item	Judgment (yes/no/unclear)	Support for judgment
External validity		
1. Was the study's target population **a close representation** of the national population in relation to relevant variables, e.g. age, sex, occupation?	Yes	According to the Japanese national census, the age and occupational distributions and the nutrient intake of residents in Hisayama have been almost identical to those of Japan as a whole during the past 50 years.
2. Was the sampling frame a **true or close representation** of the target population?	Yes	All residents of the town Hisayama aged at least 65.
3. Was some form of **random selection** used to select the sample, OR, was a census undertaken?	Yes	All residents of the town Hisayama aged at least 65 years were approached.
4. Was the likelihood of **non-response bias minimal (≥75%)?**	Yes	In 1988, a total of 837 residents aged ≥65 years participated in a screening survey (participation rate 91.8%). Similarly, a total of 1,353 residents aged ≥65 years underwent a screening survey conducted in 2002 (participation rate 83.2%).
Internal validity		
5. Were data collected **directly from the subjects** (as opposed to a proxy)?	Yes	Screening surveys were conducted among the participants of both cohorts.
6. Was an **acceptable case definition** used in the study?	Yes	Diagnosis of dementia was made based on the guidelines of the DSM-III-R.
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)**?	Yes	Neuropsychological tests comprised the Hasegawa Dementia Scale (HDS), the Hasegawa Dementia Scale - Revised (HDS-R) - Revised, or the Mini- Mental State Examination (MMSE). The HDS and HDS-R are commonly used neuropsychological tests in Japan, and their results have been shown to agree well with the MMSE.
8. Was the **same mode of data collection** used for all subjects?	Yes	Similar surveillance methods, including similar diagnostic criteria, were used in all surveys.
9. Was the **length of the incidence period** for the parameter of interest appropriate? Is the **length of the time interval to rate incidence changes** appropriate?	Yes	Follow-up screening surveys were conducted in 1992, 1998, 2005, and 2012. Each cohort was followed prospectively for 10 years.
10. Was a comprehensible and **appropriate approach** followed to **establish incidence rate** and **incidence**	Yes	The incidence of dementia and its subtypes were calculated by the person-year method with adjustment for

rate change?		age and sex by the direct method with 5-year age groups. The age- and sex-adjusted incidence was compared between the two cohorts using the Cox proportional hazards model, and the adjusted hazard ratio (HR) with its 95% confidence intervals (CI) was also estimated using the Cox model.
Summary item on the overall risk of study bias		
LOW RISK OF BIAS: Further research is very unlikely to change our confidence in the estimate.		

References

1. Hoy D, Brooks P, Woolf A, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. J Clin Epidemiol. 2012;65:934–939.

Table S1 Database for main meta-analysis of studies on trends in dementia incidence from high-income countries (n = 5).

Study	Country	Cohort	IC	IC 95%CI	IR	IR 95%CI	Study begin (year)	Duration (years)	Age (mean)	Sex (female, %)	Response rate (%)	Diagnostic criteria	Diagnostic approach
Rotterdam Study	Netherlands (Europe)	cohort 1	1.00		6.56	5.6–7.7	1990	5	71.3	59.60	78.2	DSM-III-R	clinical
		cohort 2	0.75	0.56–1.02	4.92	3.7–6.6	2000	5	67.8[a]	56.10	67.3		
IIDP	USA	cohort 1	1.00		36.0	32.0–41.0	1992	17	77.7	65.50	86.0	DSM-III-R	clinical
		cohort 2	0.40	0.30–0.50	14.0	12.0–17.0	2001	8	77.2	65.50	44.0		
PAQUID & 3C	France (Europe)	cohort 1	1.00		19.6[b]	16.4–23.5	1988	10	75.0	61.70	60.0	DSM-5	clinical
		cohort 2	0.92	0.73–1.15	20.4[b]	17.5–23.7	2000	10	74.6	61.20	39.0		
CFAS I & II	UK (Europe)	cohort 1	1.00		20.1	16.8–24.0	1992	2	75.4[a]	61.00	80.0	DSM-III-R	algorithmic
		cohort 2	0.80	0.60–1.00	17.7	15.2–20.6	2010	2	75.7[a]	56.00	56.0		
Hisayama Study	Japan	cohort 1	1.00		25.9	21.7–30.9	1988	10	74.0	61.02	91.8	DSM-III-R	clinical
		cohort 2	1.68	1.38–2.06	41.6	37.0–46.1	2002	10	73.0	57.03	83.2		

Abbreviations: 95%CI, 95% confidence interval; IC, incidence change/ratio of incidence rates of follow-up to reference period; IR, incidence rate/number of dementia events occurring during study /1,000 person years (where necessary, originally reported IR were transformed to yield comparable estimated); CFAS, Cognitive Function and Ageing Study, IIDP, Indianapolis-Ibadan Dementia Project; PAQUID & 3C, Personnes Agées Quid & Three-City; DSM, Diagnostic and Statistical Manual of Mental Disorders

[a]extrapolated from mean categorical age, weighted by number of cases

[b]estimates from L. Grasset (personal communication)

Table S2 Results of random-effect meta-regression analyses regarding incidence change in studies of sufficient methodological quality.

	Variables		**Studies**							
			n = 5[a]				n = 4[b]			
			Estimate	95%CI	SE	*p*	Estimate	95%CI	SE	*p*
Design effect	Response rate (%)		1.06	0.97-1.17	0.01	.07				
	Diagnostic criteria	DSM-IV&5 vs. DSM-III	5.77	0.27-124.60	1.40	.09	*insufficient observations*			
	Diagnostic approach	Algorithmic vs. clinical	1.58	0.20-12.62	0.26	.22				
Time effect[c]	Calendar year		0.81	0.62-1.06	0.05	.08	0.86	0.41-1.78	0.05	.23
	Length of study		1.08	0.93-1.26	0.04	.15	1.07	0.75-1.53	0.03	.26
Adjusted time effect	Calendar year		0.82	0.34-1.97	0.06	.22				
	Length of study (years)		1.10	0.62-1.98	0.05	.28				
	Age, mean (years)		0.91	0.21-4.06	0.11	.58				
	Sex, female (%)									
Adjusted time effect with country	Calendar year		*insufficient observations*				*insufficient observations*			
	Length of study (years)									
	Age, mean (years)									
	Sex, female (%)									
	Country, ref. EU	USA								
		Japan								
Publication bias	Egger's test				*t* = -1.77, *p* = .18				*t* = -0.34, *p* = .77	

Abbreviations: 95%CI, 95% confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; EU, European Union; n/a, not applicable; SE, standard error; USA, United States of America

[a]includes the following studies: Rotterdam Study, Indianapolis-Ibadan Dementia Project (IIDP), Personnes Agées Quid (PAQUID) & Three-City (3C), Cognitive Function and Ageing Study (CFAS I & II), Hisayama Study

[b]includes all studies in (a), except for the Hisayama Study

[c]includes the start year of the first observation period and an indicator of "elapsed time", i.e. the difference (in years) between the start of the first and the start of the second observation period as predictors

Table S3 Results of random-effect meta-regression analyses regarding incidence rate in studies of sufficient methodological quality.

Variables			Cohorts							
			n=10[a]				n=8[b]			
			Estimate	95%CI	SE	*p*	Estimate	95%CI	SE	*p*
Design effect	Response rate (%)		1.02	0.98–1.06	0.02	.32	1.01	0.96–1.06	0.02	.71
	Diagnostic criteria	DSM-IV&5 vs. DSM-III	1.96	0.32–12.08	1.46	.40	2.05	0.25–16.72	1.55	.40
	Diagnostic approach	Algorithmic vs. clinical	1.32	0.29–6.04	0.82	.67	1.67	0.27–10.30	1.10	.48
Time effect[c]	Calendar year		1.02	0.95–1.09	0.03	.62	1.01	0.92–1.11	0.04	.77
	Length of study		1.09	0.98–1.23	0.53	.10	1.08	0.94–1.23	0.06	.22
Adjusted time effect	Calendar year		0.98	0.92–1.03	0.22	.32	0.98	0.92–1.04	0.02	.34
	Length of study (years)		1.10	1.01–1.20	0.38	<.05	1.06	0.96–1.16	0.03	.18
	Age, mean (years)		1.30	1.08–1.56	0.09	<.05	1.27	1.06–1.51	0.07	.02
	Sex, female (%)		.082	0.68–0.99	0.06	<.05	0.89	0.73–1.09	0.06	.17
Adjusted time effect with country	Calendar year		1.00	0.91–1.09	0.03	.94	0.99	0.88–1.12	0.27	.79
	Length of study (years)		1.07	0.96–1.20	0.04	.14	1.06	0.91–1.23	0.04	.23
	Age, mean (years)		1.25	1.01–1.54	0.08	<.05	1.25	0.95–1.64	0.08	.07
	Sex, female (%)		0.92	0.65–1.31	0.10	.52	0.94	0.60–1.49	0.10	.63
	Country, ref. EU	USA	0.70	0.10–4.79	0.42	.60	0.67	0.06–7.94	0.38	.56
		Japan	1.78	0.65–4.84	0.56	.17	n/a	n/a	n/a	n/a
Publication bias	Egger's test				t = -3.08; p < .05				t = -2.16; p = .07	

Abbreviations: 95%CI, 95% confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; EU, European Union; n/a, not applicable; SE, standard error; USA, United States of America

[a]includes 2 cohorts each of the following studies: Rotterdam Study, Indianapolis-Ibadan Dementia Project (IIDP), Personnes Agées Quid (PAQUID) & Three-City (3C), Cognitive Function and Ageing Study (CFAS I & II), Hisayama Study

[b]includes cohorts from all studies from Western high-income countries as in (a), except for the cohorts of the Hisayama Study (Japan/East Asian high-income country)

[c]the model included the start year and the length of each of the observation period as predictors in the regression models

Table S4 Results of random-effect meta-regression analyses regarding incidence change across all included studies.

	Variables		Studies							
			n=7[a]				n=6[b]			
			Estimate	95%CI	SE	*p*	Estimate	95%CI	SE	*p*
Design effect	Response rate (%)		1.02	0.96–1.08	0.02	.37	0.99	0.93–1.05	0.02	.67
	Diagnostic criteria	DSM-IV&5 vs. DSM-III	1.13	0.32–3.92	0.51	.80	1.25	0.48–3.27	0.38	.51
		NINCDS-ADRDA vs. DSM-III	1.12	0.24–5.33	0.63	.85	1.95	0.51–7.51	0.83	.21
	Diagnostic approach	Algorithmic vs. clinical	1.16	0.07–5.66	0.49	.58	1.46	0.41–5.25	0.59	.41
Time effect[c]	Calendar year		1.01	0.96–1.07	0.02	.64	1.01	0.97–1.05	0.02	.64
	Length of study		1.00	0.93–1.08	0.03	.98	1.00	0.94–1.06	0.02	.90
Adjusted time effect	Calendar year		1.00	0.92–1.09	0.03	.97	1.01	0.94–1.09	0.03	.69
	Length of study (years)		1.00	0.92–1.09	0.03	.94	1.00	0.02–1.07	0.03	.88
	Age, mean (years)		0.95	0.69–1.30	0.12	.69	1.02	0.77–1.34	.10	.88
	Sex, female (%)		*insufficient observations*				*insufficient observations*			
Adjusted time effect with country	Calendar year		1.02	0.94–1.11	0.03	.52	1.02	0.94–1.11	0.03	.52
	Length of study (years)		0.98	0.90–1.07	0.03	.60	0.98	0.90–1.07	0.03	.60
	Age, mean (years)		1.12	0.75–1.68	0.14	.42	1.12	0.75–1.68	0.14	.42
	Sex, female (%)		*insufficient observations*				*insufficient observations*			
	Country, ref. EU	USA	0.64	0.20–2.08	0.24	.31	0.64	0.20–2.08	0.24	.31
		Japan	2.11	0.58–7.74	0.86	.16	n/a	n/a	n/a	n/a
Publication bias	Egger's test				t = -2.78; p = .22				t = -3.62; p < .05	
Heterogeneity	I^2				91.9%				88.2%	

Abbreviations: 95%CI, 95% confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; EU, European Union; n/a, not applicable; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; SE, standard error; USA, United States of America

[a]includes the following studies: Rotterdam Study, Indianapolis-Ibadan Dementia Project (IIDP), Personnes Agées Quid (PAQUID) & Three-City (3C), Cognitive Function and Ageing Study (CFAS I & II), Hisayama Study, Framingham Heart Study (FHS), Chicago Health and Aging Project (CHAP)

[b]includes all studies in (a), except for the Hisayama Study

[c]includes the start year of the first observation period and an indicator of "elapsed time", i.e. the difference (in years) between the start of the first and the start of the second observation period as predictors

Table S5 Results of random-effect meta-regression analyses regarding incidence rate across all included studies.

Variables			Cohorts							
			n=15[a]				n=13[b]			
			Estimate	95%CI	SE	*p*	Estimate	95%CI	SE	*p*
Design effect	Response rate (%)		1.02	0.99–1.04	0.12	.18	1.01	0.98–1.04	0.01	.48
	Diagnostic criteria	DSM-IV&5 vs. DSM-III	1.82	0.78–4.21	0.69	.15	2.21	0.94–5.21	0.82	.07
		NINCDS-ADRDA vs. DSM-III	1.25	0.30–5.20	0.80	.74	1.73	0.40–7.44	1.09	.41
	Diagnostic approach	Algorithmic vs. clinical	1.31	0.44–3.93	0.65	.60	1.67	0.54–5.09	0.01	.32
Time effect	Calendar year		0.99	0.95–1.02	0.18	.42	0.98	0.94–1.02	0.02	.28
	Length of study		1.05	0.97–1.14	0.04	.18	1.04	0.95–1.14	0.04	.32
Adjusted time effect	Calendar year		0.96	0.92–1.00	0.02	.07	0.96	0.91–1.00	0.02	.06
	Length of study (years)		1.07	0.97–1.18	0.05	.17	1.03	0.92–1.16	0.05	.54
	Age, mean (years)		1.14	0.98–1.33	0.08	.08	1.15	0.98–1.35	0.08	.08
	Sex, female (%)		0.87	0.73–1.03	0.07	.09	0.90	0.75–1.08	0.07	.23
Adjusted time effect with country	Calendar year		0.97	0.92–1.02	0.02	.16	0.96	0.92–1.01	0.02	.12
	Length of study (years)		1.03	0.92–1.15	0.05	.61	1.02	0.90–1.14	0.05	.77
	Age, mean (years)		1.12	0.96–1.31	0.08	.12	1.13	0.96–1.32	0.08	.11
	Sex, female (%)		0.91	0.76–1.08	0.07	.25	0.92	0.77–1.11	0.07	.33
	Country, ref. EU	USA	1.51	0.76–3.03	0.45	.21	1.49	0.73–3.02	0.45	.22
		Japan	2.05	0.71–5.96	0.95	.16	n/a	n/a	n/a	n/a
Publication bias	Egger's test				t = -8.69; p = .14				t = -5.46; p = .38	
Heterogeneity	I^2				97.6%				97.2%	

95%CI, 95% confidence interval; DSM, Diagnostic and Statistical Manual of Mental Disorders; EU, European Union; n/a, not applicable; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; SE, standard error; USA, United States of America

[a]includes the following studies: Rotterdam Study, Indianapolis-Ibadan Dementia Project (IIDP), Personnes Agées Quid (PAQUID) & Three-City (3C), Cognitive Function and Ageing Study (CFAS I & II), Hisayama Study, Framingham Heart Study (FHS), Chicago Health and Aging Project (CHAP)

[b]includes all studies in (a), except for the Hisayama Study

[c]the model included the start year and the length of each of the observation period as predictors in the regression models