ORIGINAL RESEARCH

Epidemiology of dementia: prevalence and incidence estimates using validated electronic health records from primary care

Anna Ponjoan[1–3]
Josep Garre-Olmo[2]
Jordi Blanch[1]
Ester Fages[1]
Lia Alves-Cabratosa[1]
Ruth Martí-Lluch[1–3]
Marc Comas-Cufí[1]
Dídac Parramon[1,4]
María Garcia-Gil[1]
Rafel Ramos[1,5]

[1]Vascular Health Research Group (ISV-Girona), Jordi Gol Institute for Primary Care Research (IDIAPJGol), Barcelona, Catalonia, Spain; [2]Girona Biomedical Research Institute (IDIBGI), Girona, Catalonia, Spain; [3]Santa Clara Primary Care Health Center, Catalan Health Institute, Girona, Catalonia, Spain; [4]Autonomous University of Barcelona, Bellaterra (Cerdanyola del Vallès), Catalonia, Spain; [5]Department of Medical Sciences, School of Medicine, Campus Salut, University of Girona, Girona, Catalonia, Spain

Correspondence: Rafel Ramos
Vascular Health Research Group (ISV-Girona) IDIAPJGol, c/ Maluquer Salvador, 11 Baixos, Girona 17002, Catalonia, Spain
Tel +34 97 248 7968
Email rramos.girona.ics@gencat.cat

Purpose: Updated estimates of incidence and prevalence of dementia are crucial to ensure adequate public health policy. However, most of the epidemiological studies in the population in Spain were conducted before 2010. This study assessed the validity of dementia diagnoses recorded in electronic health records contained in a large primary-care database to determine if they could be used for research purposes. Then, to update the epidemiology of dementia in Catalonia (Spain), we estimated crude and standardized prevalence and incidence rates of dementia in Catalonia in 2016.

Methods: The System for the Development of Research in Primary Care (SIDIAP) database contains anonymized information for >80% of the Catalan population. Validity of dementia codes in SIDIAP was assessed in patients at least 40 years old by asking general practitioners for additional evidence to support the diagnosis. Crude and standardized incidence and prevalence (95% CI) in people aged ≥65 years were estimated assuming a Poisson distribution.

Results: The positive predictive value of dementia diagnoses recorded in SIDIAP was estimated as 91.0% (95% CI 87.5%–94.5%). Age- and sex-standardized incidence and prevalence of dementia were 8.6/1,000 person-years (95% CI 8.0–9.3) and 5.1% (95% CI 4.5%–5.7%), respectively.

Conclusion: SIDIAP contains valid dementia records. We observed incidence and prevalence estimations similar to recent face-to-face studies conducted in Spain and higher than studies using electronic health data from other European populations.

Keywords: family physician, accuracy, quality, positive predictive value, electronic medical record, real-world data

Introduction

Dementia affects about 47 million people worldwide and is related to dependence, poor quality of life, institutionalization, and mortality.[1] Age is the main risk factor for dementia, and the world population is aging; therefore, the number of people affected by dementia is expected to increase to 131 million in 2050, resulting in huge social and economic costs.[1] Population-based epidemiological estimates of its prevalence and incidence is crucial to plan and budget health services and evaluate its economic burden. Therefore, robust and up-to-date estimates are needed to support the creation of useful public policies on dementia.[2] However, updated estimates of prevalence and incidence in the European population are scarce,[1] eg, in Spain, most of the studies collected data in the 1990s or 2000s, and to our knowledge, only four used data collected during the past 10 years.[3–6] High economic costs of door-to-door surveys or questionnaire-based studies may contribute to the difficulty of updating estimates of prevalence and incidence of dementia.

Electronic health records emerged as a new opportunity to study the epidemiology of dementia: prevalence and incidence-rate estimates can be updated at low cost using data routinely collected in health care settings. Indeed, the use of electronic medical databases to assess the epidemiology of dementia has been increasing at the European level[7] and in many developed countries, such as Canada[8] and the Netherlands.[9] However, the accuracy of dementia records in routinely collected health care data sets is always a concern, and cohorts should ideally conduct their own setting-specific validation.[10] This concern is especially apparent for electronic health records collected in primary care, because accuracy of dementia diagnoses recorded by general practitioners has been considered low.[8,11] Nevertheless, a recent systematic review reported high positive predictive values (PPVs) (0.83 or higher) when identifying dementia cases in primary-care data, indicating that dementia codes in those primary-care databases were sufficiently accurate to be used for research purposes.[10]

We thus used data from a large primary-care database to describe the epidemiology of dementia in Catalonia, southern Europe. First, we assessed the validity of dementia diagnoses recorded in the database to determine if they could be used for research purposes. Then, we estimated crude and standardized prevalence and incidence rates of dementia in Catalonia in 2016, in order to update the epidemiological estimates.

Methods

This population-based study was approved by the Clinical Research Ethics Committee, Primary Care Research Institute – IDIAP Jordi Gol. Data were obtained from the System for the Development of Research in Primary Care (SIDIAP) database, which contains routine records of consultations from ~275 general-care practices belonging to the National Health Service of Catalonia (Institut Català de la Salut). SIDIAP contains anonymized longitudinal medical records containing data on demographics, symptoms, diagnoses, and prescriptions for about 6 million people (>80% of the Catalan population, 15% of the Spanish population).[12,13] We defined dementia diagnoses using the following codes from the International Classification of Diseases, Tenth revision (ICD10): dementia in Alzheimer's disease (F00), vascular dementia (F01), dementia in other diseases classified elsewhere (F02.8), unspecified dementia (F03), Alzheimer's disease (G30), Lewy body disease (G31.8), or frontotemporal dementia (G31.0). SIDIAP is linked with the pharmacy-invoicing database provided by the National Health Service of Catalonia and medications are recorded using the Anatomical

Therapeutic Chemical (ATC) codes. We considered as treated those patients who had a prescription or billing for anticholinesterases (N06DA) or memantine (N06D × 01). We used outcomes recorded in SIDIAP to describe dementia assessment in the Catalan primary-care system, because clinical practice might influence the identification of dementia cases, and this would affect estimates of prevalence and incidence rates.[10,14] In particular, we identified the number of persons who were institutionalized, attended by the home-care program, had a record of cognitive assessment, or had a record of functional ability and independence assessment. Cognitive assessment was defined as at least one Mini-Mental Status Examination[15] or Pfeiffer test[16] (Spanish versions) score ≤23 or ≥3, respectively, considered to suggest cognitive impairment. Assessment of functional ability and independence was defined as at least one recorded score on the Barthel Index,[17] Lawton Scale,[18] Katz Index,[19] or Blessed Dementia Rating Scale.[20] Low functional ability or dependence was considered when the score was <90 on the Barthel Index, <3 in men and <5 in women on the Lawton Scale, ≥2 on the Katz Index, or ≥4 on the Blessed Dementia Rating Scale.

Validation of dementia diagnoses

To assess accuracy of dementia diagnoses recorded in SIDIAP, we emailed a short survey to general practitioners asking about the current status of the diagnosis (Figure S1), a method used previously in validations of dementia diagnosis.[21] We invited general practitioners of the Agency of Clinical Research Management in Primary Care (AGICAP), associated with the IDIAP Jordi Gol to participate in the validation study. AGICAP general practitioners received training and had experience in recruitment of patients in clinical trials and in review of diagnoses recorded in the electronic medical history. We emailed an online survey to general practitioners who agreed to participate, asking for further evidence to confirm the diagnosis of all their patients aged at least 40 years who had a dementia code recorded in SIDIAP (ICD10). The survey remained available online for 1 month for general practitioners to evaluate dementia diagnoses. In order to enhance the response rate, we offered a small monetary incentive and sent one follow-up email to nonresponding general practitioners to remind them to complete the questionnaire. The online survey examined whether or not the diagnosis was made by a specialist, whether or not the diagnosis fulfilled the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) IV or ICD10 diagnosis criteria, which symptoms were observed, subtype and severity of dementia, and if the diagnosis was considered to be incor-

rect, evolved, or uncertain due to incomplete information (Figure S1). When applicable, additional information about the score achieved in cognitive and functional tests was requested (Figure S1). We defined patients as true cases if they were diagnosed by a specialist (eg, neurologist or psychiatrist), by cognitive or functional tests, or by clinical opinion of the general practitioner and DSM-IV or ICD10 diagnostic criteria were fulfilled (Figure 1). Patients were defined as false cases when the diagnosis was considered by the general practitioner not to fulfill DSM-IV or ICD10 diagnostic criteria, was based on inconsistent evidence, had evolved, or was incorrect (Figure 1).

Incidence and prevalence

The study population included patients aged at least 65 years who were registered in SIDIAP. We defined a case as any patient having an ICD10 diagnosis of dementia or an ATC code of pharmacological treatment related to dementia recorded in SIDIAP. In Catalonia, the prescription of antidementia drugs is requested by the general practitioner, but until 2014 required approval from the advisory board, a group of experts who evaluated all patients with dementia recommended for pharmacological treatment. Since 2014,

approval has been determined by a geriatrician, psychiatrist, or neurologist. Since all treated patients had been evaluated previously by specialists, we considered them true cases. To estimate the incidence rate, persons who were alive in 2016 according to SIDIAP records were included. Incident cases were defined as dementia patients diagnosed or treated during 2016. Patients diagnosed with dementia or treated with antidementia drugs before January 1, 2016 were excluded. To estimate dementia prevalence, persons who had SIDIAP records and were alive as of December 31, 2016 were included. Patients who moved away or died before that date were excluded. Prevalent cases were defined as patients diagnosed or treated as of December 31, 2016.

Statistical analyses

We estimated PPV and 95% CI of dementia diagnosis using data from the survey of general practitioners. PPV is an indicator of the probability that a person with a record of dementia diagnosis in SIDIAP truly had the disease. A higher PPV indicated better accuracy of the dementia records in SIDIAP. PPV was defined as the number of confirmed diagnoses of dementia (numerator) against all the diagnoses of dementia evaluated by general practitioners using the questionnaire
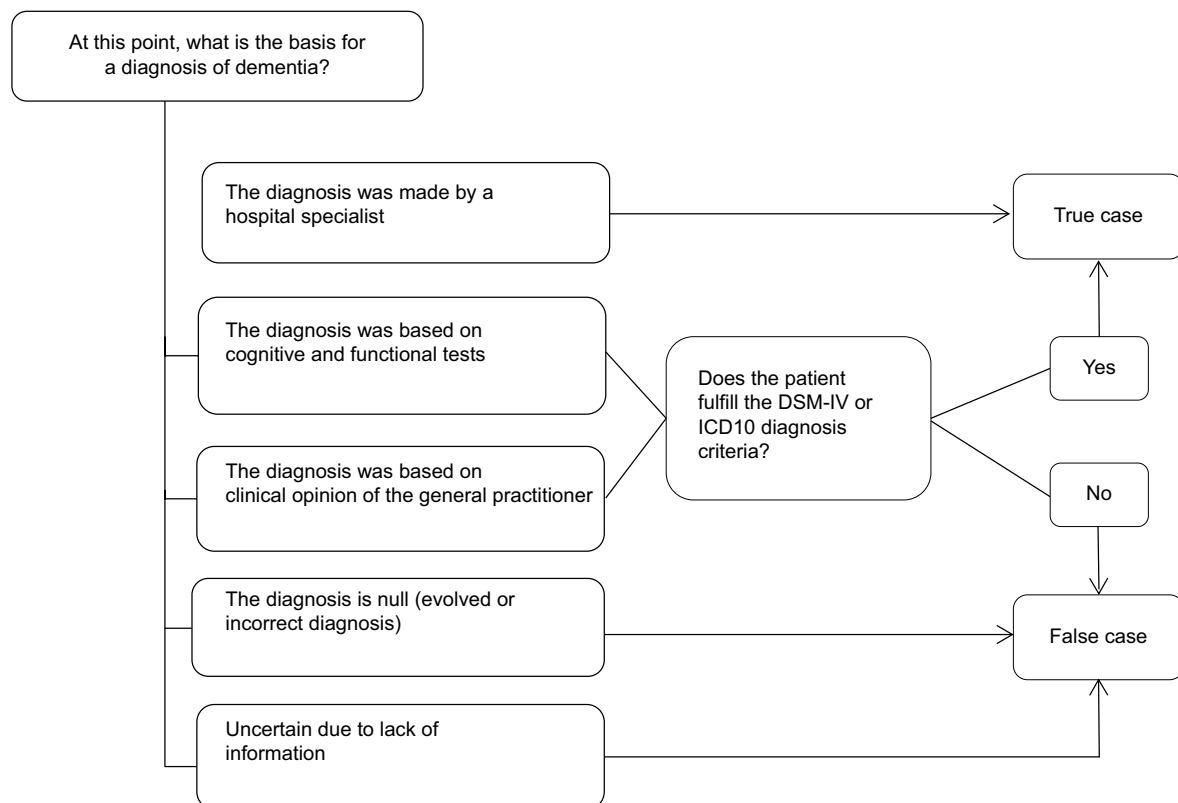

Figure 1 Definition of true and false cases of dementia.
Abbreviations: DSM, *Diagnostic and Statistical Manual of Mental Disorders*; ICD10, International Classification of Diseases, Tenth revision.

(denominator). Results were stratified by sex and age-group. We also performed sensitivity analyses to explore the robustness of our results. By extrapolating figures to the worst-case scenario, we replicated the analyses considering both evaluated and unevaluated diagnoses of dementia. PPV was replicated, including in the numerator all diagnoses of dementia confirmed by the general practitioners or with evidence of treatment with antidementia drugs, and in the denominator all recorded diagnoses of dementia included in the validation study (evaluated or not by general practitioners).

To describe the study population and primary-care settings, we used percentages for categorical variables and means (SD) or medians (IQR) for continuous variables. We estimated overall crude prevalence and incidence rates, and 95% CI was calculated assuming a Poisson distribution. We stratified the prevalence and incidence estimates by sex and age-group. We used the direct method to estimate age- and sex-standardized prevalence and incidence of dementia. The 2013 revision of the European standard population served as the basis for adjustment (EU27/European Free Trade Association standard population:[22] weights of 0.28, 0.26, 0.21, 0.13, 0.08, and 0.05 for age-groups 65–69, 70–74, 75–79, 80–84, 85–89, and ≥90 years, and 0.5 for both men and women). We used the official 2015 population estimate of 11,045,521 citizens in Spain aged ≥65 years (http://www.ine.es).

Finally, we calculated age-standardized rates for women and men and sex-standardized rates for age-groups separately. We used the likelihood-ratio test to examine the age–sex interaction. All analyses were performed using R software version 3.5.2.[23]

Results
Validation of dementia diagnoses
Surveys were sent to 42 general practitioners, of which 29 participated in the validation study (response rate 69%). These 29 general practitioners had records of 374 patients with a dementia diagnosis, of which 256 diagnoses were evaluated during the 1-month validation period. Among the 118 unevaluated dementia diagnoses, 58 of the patients were treated with antidementia drugs and 60 were untreated. General practitioners evaluated 256 dementia diagnoses and considered 233 true cases: 209 patients diagnosed by a specialist and 24 patients who fulfilled the DSM-IV or ICD10 diagnostic criteria. We identified 23 false cases: 21 considered evolved, incorrect, or uncertain diagnoses and two not fulfilling DSM-IV or ICD10 diagnostic criteria. Therefore, the PPV of dementia codes registered in SIDIAP was 91.0% (95% CI 87.5%–94.5%). Results stratified by age and sex provided similar results: PPV

was estimated as 89.8% (95% CI 85.2%–94.4%) for women (n=166), 93.3% (95% CI 88.1%–98.5%) for men (n=90), 90.3% (95% CI 85.6%–95.0%) for persons aged <80 years (n=154), and 92.1% (95% CI 86.9%–97.3%) for persons at least 80 years old (n=102).

General practitioners evaluated 128 diagnoses of patients who had both a diagnosis and pharmacological treatment for dementia recorded in SIDIAP, and 99% of these were confirmed as true cases. We replicated the PPV estimates, adding the unevaluated cases as true cases if the diagnosed patient had been treated with antidementia drugs or as false cases if they had not been treated. In this worst-case scenario, the PPV estimate included 291 true cases in the numerator (233 dementia diagnoses evaluated and confirmed by general practitioners, and 58 unevaluated dementia diagnoses in patients being treated for the disease) and 374 dementia diagnoses in the denominator (256 evaluated by general practitioners and 118 unevaluated). The resulting PPV was 77.8% (95% CI 73.6%–82.0%).

Incidence and prevalence
For 2016, we obtained data for 1,035,046 persons, mainly women (56.2%), from urban areas (80.9%) and 75.7 (7.9) years old on average. We obtained 1,008,755 person-years and detected 9,596 incident cases, of which 8,553 had been diagnosed and 1,043 undiagnosed but treated with antidementia drugs. Table S1 provides further details of the population studied to estimate incidence rates. The crude incidence of dementia was estimated at 9.5/1,000 person-years (95% CI 9.3–9.7), sex-standardized incidence at 9.3/1,000 person-years (95% CI 9.0–9.6), age-standardized incidence at 8.8/1,000 person-years (95% CI 8.4–9.2), and age- and sex-standardized incidence at 8.6/1,000 person-years (95% CI 8.0–9.3). The incidence rate was higher among women than men, and increased with age: it was about 25 times higher in people older than 90 years than in people aged 65–69 years (Table 1). We found a significant age–sex interaction ($P<0.001$), and age-related increase in incidence was more pronounced in women than in men (Figure 2, Table S2).

To estimate dementia prevalence, we obtained data on 1,048,956 persons recorded in SIDIAP on December 31, 2016. Participants were mainly women (57.2%) and predominantly from urban areas (80.9%). Mean (SD) age was 75.9 (7.9) years. Table S3 shows the medical conditions of the study population used to estimate dementia prevalence. We found 63,117 prevalent cases, of which 55,951 had been diagnosed and 7,166 undiagnosed but treated with antidementia drugs. The median duration of dementia in prevalent

Table 1 Crude and standardized incidence rates of dementia (per 1,000 person-years) in people aged ≥65 years in Catalonia (Spain) in 2016, stratified by sex and age

Groups	Cases	Person/years	Crude incidence (95% CI)	Standardized incidence (95% CI)
Whole population	9,596	1,008,755	9.5 (9.3–9.7)	8.6 (8.0–9.3)[a]
Sex				
Women	6,218	567,836	11.0 (10.7–11.2)	9.4 (8.9–10.0)[b]
Men	3,378	440,919	7.7 (7.4–7.9)	7.8 (7.2–8.5)[b]
Age				
65–69 years	330	282,876	1.2 (1.0–1.3)	1.2 (1.0–1.4)[c]
70–74 years	920	240,706	3.8 (3.6–4.1)	3.8 (3.5–4.2)[c]
75–79 years	1,574	173,703	9.1 (8.6–9.5)	9.0 (8.4–9.6)[c]
80–84 years	2,667	159,296	16.7 (16.1–17.4)	16.4 (15.5–17.3)[c]
85–89 years	2,490	97,631	25.5 (24.5–26.5)	24.9 (23.5–26.4)[c]
≥90 years	1,615	54,543	29.6 (28.2–31.1)	28.6 (26.5–30.8)[c]

Note: [a]Age- and sex-standardized; [b]age-standardized; [c]sex-standardized.

cases was 3.77 (IQR 1.67–6.50) years. Table 2 describes the dementia assessment in the Catalan primary settings. We estimated crude prevalence as 6.0% (95% CI 6.0%–6.1%), sex-standardized prevalence 5.8% (95% CI 5.7%–5.9%), age-standardized prevalence 5.3% (95% CI 5.0%–5.6%), and both sex- and age-standardized prevalence 5.1% (95% CI 4.5%–5.7%). Dementia prevalence was higher in women than men and increased with age (Table 3). The age-related increase in prevalence was significantly higher in women than in men, with a *P*-value <0.001 for the age–sex interaction (Figure 2, Table S4).

Discussion

This large-scale nationwide study of >1 million cases provides a comprehensive update of the epidemiology of dementia in southern Europe. Our findings highlighted that records of dementia diagnoses routinely collected in primary care and registered in SIDIAP are accurate enough to be used for research purposes. Moreover, we provided a detailed update about crude and standardized prevalence and incidence rates, which may help to improve policies related to dementia not only in primary-care settings but also in secondary care.

Validation of dementia diagnoses

We found a 91% PPV of diagnoses of dementia recorded in SIDIAP. Sensitivity analyses showed that in the worst scenario, the PPV would be 77.8%. Our PPV estimate was slightly higher than values observed in previous studies that assessed the quality of routinely collected data on dementia in primary-care services for research purposes.[10,24,25] When routinely collected electronic health records are applied to research, avoiding false-positive diagnoses could be more important than avoiding false negatives. In longitudinal studies, for example, false positives can dilute observed

effects and reduce statistical power.[26] Our results highlighted that false positives on dementia cases are scarce in SIDIAP and that patients having a record in SIDIAP of a diagnosis of dementia are very likely to have the disease. Therefore, the SIDIAP database is valid for epidemiological research on dementia.

Incidence and prevalence

We estimated an age- and sex-standardized incidence of 8.6/1,000 person-years (95% CI 8.0–9.3), similar to the adjusted incidence estimated by the NEDICES study on the Spanish population aged ≥65 years.[27] Other studies conducted in Spain have reported disparate estimates of dementia incidence.[28,29] However, differences in study populations and methodologies limit comparability with our findings. When comparing incidence of dementia according to age-group, our estimations were similar to those reported by Bermejo-Pareja et al[27] and Lobo et al[28] in all age-groups except the oldest old (aged at least 85 years), in which our values were lower than the face-to-face studies cited. This disparity might be partially due to some level of underrecording in SIDIAP, mainly concentrated in the oldest-old patients, in line with previous studies.[30] Some general practitioners may consider memory difficulties part of the normal aging process, rather than a disability that requires specialist care and support,[31] while others might be reluctant to use dementia codes in this age-group because they consider the diagnosis and treatment of dementia useless.[8,31]

We estimated the age- and sex-standardized prevalence of dementia as 5.1% (95% CI 4.5%–5.7%), which overlapped with the standardized estimations reported by the the most recent face-to-face studies in the Spanish population at least 65 years old.[4,5] Previous face-to-face studies – mostly conducted before 2010 – reported crude prevalence estimates, ranging from 5.5% to 14.9% in Spanish people aged ≥65
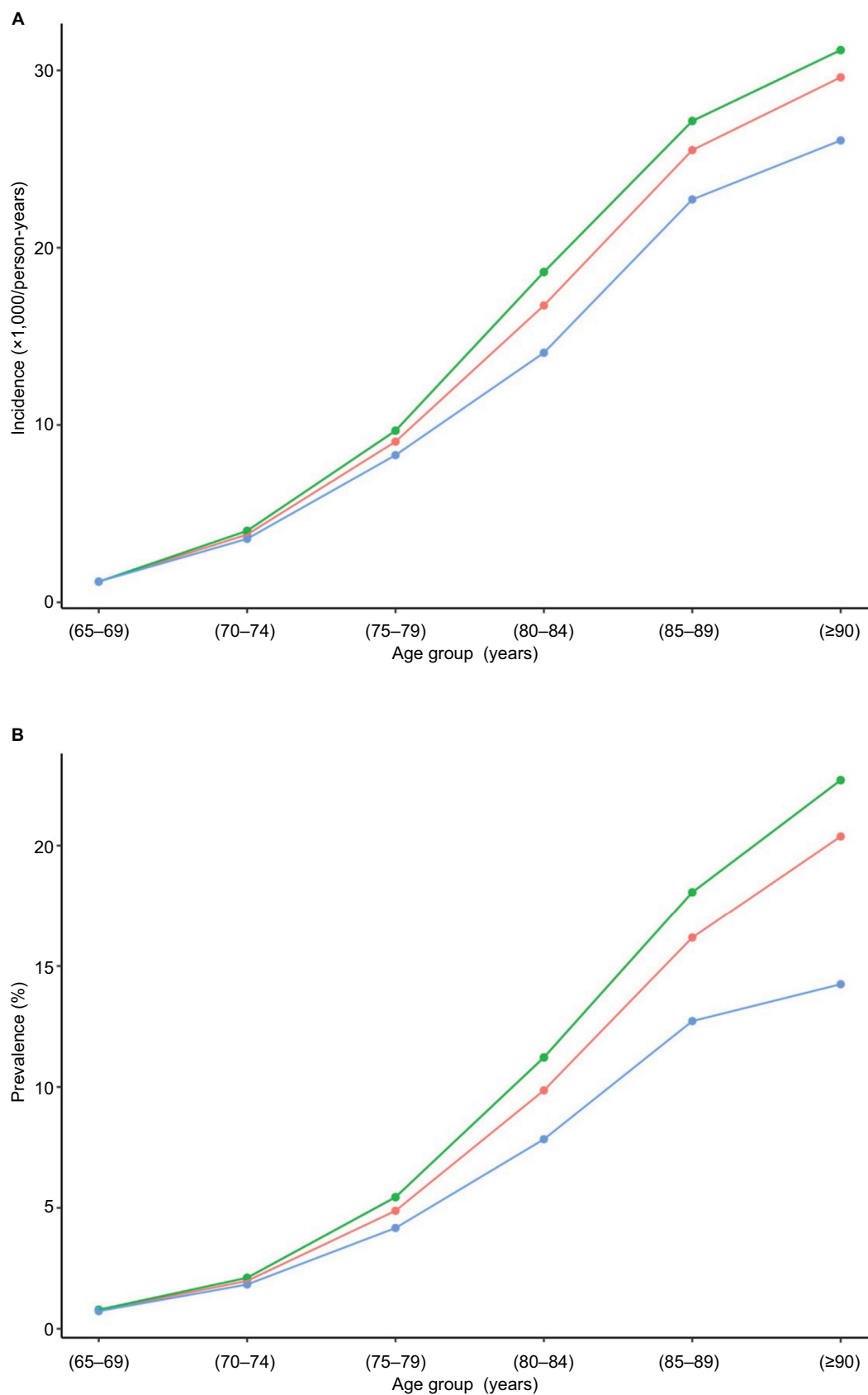
A

B

Figure 2 Incidence rate (**A**) and prevalence (**B**) of dementia.
Note: According to sex (green line, women; blue line, men; red line, overall population) by age-group in Catalan population aged at least 65 years.

Table 2 Description of the characteristics of the primary care settings in Catalonia (Spain) in relation to dementia assessment

Dementia assessment	Study population (n=1,048,956), n (%)	Dementia-free patients (n=985,839), n (%)	Patients with dementia (n=63,117), n (%)
Patients at home-care program	67,441 (6.4)	50,944 (5.2)	16,497 (26.1)
Institutionalized patients	41,825 (4.0)	23,288 (2.4)	18,537 (29.4)
Cognitive abilities			
Persons assessed	268,229 (25.6)	220,506 (22.4)	47,723 (75.6)
Persons impaired	76,410 (7.3)	40,787 (4.1)	35,623 (56.4)
Functional ability and independence			
Persons assessed	242,251 (23.1)	201,203 (20.4)	41,048 (65.0)
Persons impaired	121,090 (11.5)	87,698 (8.9)	33,392 (52.9)

Table 3 Crude and adjusted prevalence of dementia in people aged ≥65 years in Catalonia (Spain) in 2016, stratified by sex and age

Groups	Cases	Population	Crude prevalence (95% CI)	Standardized prevalence (95% CI)
Whole population	63,117	1,048,956	6.0 (6.0–6.1)	5.1 (4.5–5.7)[a]
Sex				
Women	43,995	599,942	7.3 (7.3–7.4)	5.9 (5.5–6.3)[b]
Men	19,122	449,014	4.3 (4.2–4.3)	4.2 (3.8–4.7)[b]
Age				
65–69 years	2,151	282,815	0.8 (0.7–0.8)	0.8 (0.7-0.8)[c]
70–74 years	4,812	242,888	2.0 (1.9–2.0)	2.0 (1.9-2.1)[c]
75–79 years	8,750	179,329	4.9 (4.8–5.0)	4.8 (4.6-5.0)[c]
80–84 years	16,906	171,494	9.9 (9.7–10.0)	9.5 (9.2–9.8)[c]
85–89 years	17,884	110,530	16.2 (16.0–16.4)	15.4 (14.9–15.8)[c]
≥90 years	12,614	61,900	20.4 (20.1–20.7)	18.5 (17.8–19.2)[c]

Note: [a]Age-and sex-standardized; [b]age-standardized; [c]sex-standardized.

years.[32–35] At the European level, age- and sex-standardized prevalence of dementia was estimated at 6.4% by a collaborative study of population-based cohorts conducted in the 1990s,[36] 7.3% by a meta-analysis that integrated studies conducted from 1980 until 2004,[37] and 7.1% based on a meta-analysis on data obtained mainly during the 1990s and 2000s.[38] More recent national estimates have been 6.5% in the UK,[39] 3.65% in Portugal,[40] or 5.0% in Greece.[41] Our prevalence estimation fell in the lower range of published results from the 1990s and early 2000s in face-to-face studies in Spain and in the European population, and were in line with the most recent estimations. The prevalence of dementia has been reported to be decreasing in high-income countries,[42] the USA,[43] the UK,[39] and Spain.[44] However, comparisons between studies are difficult to make, because of methodological differences. Therefore, it is plausible that our 2016 estimates reflect a reduction in dementia cases since the 1990s or 2000s, but we cannot reject other explanations. Population characteristics could be a factor. For example, most of the population in Catalonia is bilingual in Catalan and Spanish. Bilingualism may contribute to cognitive reserve and protect against the onset of dementia,[45] and could have affected our estimates. SIDIAP does not provide data on languages spoken by patients.

Electronic health records

In a comparison with other studies based on electronic health records, our findings provided new insights about the epidemiology of dementia. Our results were similar to those reported by a study that used diagnoses and drug-prescription data from primary-care and pharmacy electronic databases to estimate unadjusted prevalence as 5.91% (95% CI 5.85%–5.95%) in persons at least 65 years old living in Madrid in 2011.[6] However, our estimates were higher than previous studies in other European populations. A Danish study estimated the prevalence of dementia at 3.8% using electronic health records from secondary-care settings.[46] Another study, by Perera et al, compiled data from six electronic health-record systems, including both primary and secondary care from different European countries, to estimate prevalence and incidence rates of dementia.[7] They reported lower incidence and prevalence estimates in all age-groups than the ones we observed. Such differences remained for estimates restricted to data from primary-care settings,

which contained fewer dementia cases than SIDIAP. These differences might be partially explained by methodological issues, demographic aspects, and characteristics of primary-care settings. Our methodology applied a mixed definition of dementia that included not only diagnosis but also prescription or billing of antidementia drugs. This broader definition might capture more dementia cases than that used by Perera et al, who identified cases using only diagnostic codes or free text.[7] Demographic aspects might also contribute to explaining differences between primary-care databases from different countries. Life expectancy in Spain is the longest in the EU (83.3 years),[47] and the sex gap in life expectancy favors women to a higher extent than in other countries, such as Denmark, the Netherlands, or the UK.[47] Since dementia prevalence is higher in women and increases with age, demographic characteristics of the population might facilitate a higher number of dementia cases in Spain than in other European countries. Primary-care characteristics might also contribute to explaining disparities between studies on dementia epidemiology based on electronic health records from primary care. Catalan primary-care settings from which SIDIAP obtained data might be highly sensitive in diagnosing dementia cases. On one hand, cognitive tests are administered widely to identify dementia cases in the Catalan primary-care system: in SIDIAP, >22% of patients with no dementia code in their electronic health history had taken a cognitive test. Patients with memory complaints reported by themselves, their relatives, or general practitioners are asked to take a cognitive test. Patients included in specific programs for home care or institutionalized due to chronic complex disease may also be asked to complete a cognitive test, even if they have no memory or cognitive complaints. The huge number of cognitive tests administered in Catalan primary care might facilitate the identification of new dementia cases.[14] On the other hand, we found that most patients with a dementia code in SIDIAP had been diagnosed by a specialist or prescribed antidementia drugs by specialists. These findings suggest a good coordination between primary and secondary settings provided by the Catalan Health Institute, strengthened by such initiatives as the Computerized Support System for the Diagnosis of Dementia in Primary Care (SISDDAP [Catalan acronym]), an asynchronous telemedicine program involving both primary- and specialized-care professionals that establishes protocols for the screening and diagnosis of dementia.[48] Appropriate communication and coordination between primary and secondary care might facilitate the identification of new dementia cases and the recording of diagnoses in the electronic health record. Our findings are

in line with previous literature reporting that settings with active dementia-diagnosis centers are likely to see a higher proportion of community cases in health care databases.[7] Our findings also suggested that the high level of underdiagnosis of dementia reported in several primary-care systems[31,49] might not be true for other primary-care settings. General practitioner concern and training, as well as programs and policies designed to better diagnose dementia cases, might help to reduce underdiagnosis in primary care. Therefore, databases containing routinely collected data from primary care might be a powerful tool to study dementia epidemiology if data have been recorded in a setting with active management of dementia cases and validated as sufficiently accurate for research purposes.

Limitations and strengths

We acknowledge several limitations. First, the PPV could have been overestimated, because the validation study was constrained to general practitioners from the AGICAP network, who are regularly involved in clinical trials and thus could be more prone to register diagnoses accurately in electronic medical records. Second, as not all dementia diagnoses were evaluated by general practitioners, we cannot discard a potential bias that might have resulted in an overestimation of the PPV. Therefore, we replicated the PPV estimate considering the worst-case scenario, in which data from both evaluated and unevaluated dementia diagnoses were used. This worst-case scenario PPV estimate would indicate the minimum value of the PPV, ie, if bias had occurred and the PPV value had been overestimated, the PPV would not be lower than the worst-case scenario estimate. Third, the validation study did not provide sensitivity estimates of dementia diagnoses recorded in SIDIAP: it was cost-prohibitive to search for cases in persons with no recorded dementia diagnosis in SIDIAP. However, sensitivity has been reported to be less crucial than PPV in evaluating the accuracy of routinely collected health data, because the effects of false negatives would be diluted among the large control population.[10] Moreover, our prevalence estimates were quite similar to those reported by clinical studies in Spain during the last decade, and a decreasing trend in prevalence figures has been observed in high-income countries.[42] All this suggests that the recording of dementia diagnoses in SIDIAP is satisfactory. Fourth, we did not use free text to improve identification of persons with dementia. However, we used a mixed definition of dementia cases that included diagnosis and prescription or billing codes involving not only data from the general practitioner but also from the pharmacy-invoicing database. Using two different

data sources minimized the possible effect of underrecording of dementia diagnoses in SIDIAP. Finally, a lower education level has been associated with a greater risk of dementia.[50] We could not stratify the results according to literacy, since this information is unavailable in SIDIAP.

Our study has several strengths. First, it included about 10,000 incident cases in 1 year and 63,000 prevalent cases, while most epidemiological studies on dementia conducted in Western Europe had smaller samples.[1] Second, our study did not depend on participants' response rate, and thus was not affected by selection bias due to lack of response, which can be common among people with dementia due to reluctance of relatives to participate or difficulties in completing the questionnaires.[1] Finally, this was a nationwide study that included >1 million people in Catalonia, a 32,000 km^2 region of southern Europe. Therefore, this study provides better representation of the Spanish population than previous studies conducted in a single city[5,34,44,51] or a limited area. The nationwide approach may facilitate the design and implementation of new dementia policies in national health systems.

Conclusion

Dementia diagnoses recorded in the SIDIAP database are valid to be used for research purposes. Age- and sex-standardized prevalence and incidence rates of dementia in 2016 were higher for women and the oldest-old population. Our figures were in line with the most recent studies conducted in Spain and higher than those provided by other primary-care databases from other European countries. These findings may be helpful to plan dementia policies not only in primary-care settings but also in secondary care and in the overall public health system.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Martin Prince A, Wimo A, Guerchet M. World Alzheimer report 2015 the global impact of dementia an analysis of prevalence, incidence, cost and trends; 2015. Available from: https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf. Accessed June 14, 2017.
2. Wu YT, Fratiglioni L, Matthews FE, et al. Dementia in Western Europe: epidemiological evidence and implications for policy making. *Lancet Neurol*. 2016;15(1):116–124.
3. Virués-Ortega J, de Pedro-Cuesta J, Vega S, et al. Prevalence and European comparison of dementia in a ≥75-year-old composite population in Spain. *Acta Neurol Scand*. 2011;123(5):316–324.
4. Rodríguez-Sánchez E, Mora-Simón S, Patino-Alonso MC, et al. Prevalence of cognitive impairment in individuals aged over 65 in an urban area: DERIVA study. *BMC Neurol*. 2011;11(1):147.
5. Tola-Arribas MA, Yugueros MI, Garea MJ, et al. Prevalence of dementia and subtypes in Valladolid, northwestern Spain: the DEMINVALL study. *PLoS One*. 2013;8(10):e77688.
6. de Hoyos-Alonso MC, Bonis J, Tapias-Merino E, Castell MV, Otero A. Estimación de la prevalencia de demencia a partir del análisis de bases de datos sobre uso de fármacos. La situación en La Comunidad de Madrid (España). *Neurología*. 2016;31(1):1–8.
7. Perera G, Pedersen L, Ansel D. Dementia prevalence and incidence in a Federation of European electronic health record databases—the European medical informatics framework resource. *Alzheimers Dement*. 2018;14(2):130–139.
8. Drummond N, Birtwhistle R, Williamson T, Khan S, Garies S, Molnar F. Prevalence and management of dementia in primary care practices with electronic medical records: a report from the Canadian primary care sentinel surveillance network. *CMAJ Open*. 2016;4(2): E177–E184.
9. van Bussel 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):e1002235.
10. Wilkinson T, Ly A, Schnier C. Identifying dementia cases with routinely collected health data: a systematic review. *Alzheimers Dement*. 2018;14(8):1038–1051.
11. Zunzunegui Pastor MV, Ser T del, Rodríguez Laso A, García Yébenes MJ, Domingo J, Otero Puime A. Demencia no detectada y utilización de los servicios sanitarios: implicaciones para la atención primaria TT - Non-detected dementia and use of the health services: implications for primary care. *Aten Primaria*. 2003;31(9):581–586.
12. Bolíbar B, Fina Avilés F, Morros R, et al. [SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research]. *Med Clin (Barc)*. 2012;138(14):617–621. Spanish
13. García-Gil MM, Hermosilla E, Prieto-Alhambra D, et al. Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). *Inform Prim Care*. 2011;19(3):135–145.
14. Eichler T, Thyrian JR, Hertel J, et al. Rates of formal diagnosis of dementia in primary care: the effect of screening. *Alzheimers Dement (Amst)*. 2015;1(1):87–93.
15. Loob A, Saz P, Marcos G, et al. [Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental status examination) in the general geriatric population]. *Med Clin (Barc)*. 1999;112(20):767–774.
16. Martínez de la Iglesia J, Dueñas Herrero R, Onís Vilches MC, Aguado Taberné C, Albert Colomer C, Luque Luque R. [Spanish language adaptation and validation of the Pfeiffer's questionnaire (SPMSQ) to detect cognitive deterioration in people over 65 years of age]. *Med Clin (Barc)*. 2001;117(4):129–134.

17. Baztán J, Pérez del Molino J, Alarcon T, San Cristóbal E, Izquierdo G, Manzarbeitia I. Indice de Barhtel: instrumento válido para la valoración funcional de pacientes con enfermedad cerebrovascular [Barthel index: a valid instrument for the functional assessment of patients with cerbrovascular disease]. *Rev Esp Geriatr Gerontol*. 1993;28:32–40. Spanish.

18. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3):179–186.

19. Alvarez Solar M, At DAR, Gurpegui B. Capacidad funcional de pacientes mayores de 65 años, según El Indice de Katz [Functional ability of patients aged 65 years or older according to the Katz index]. *Aten Primaria*. 1992;10:812–815.

20. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry*. 1968;114(512):797–811.

21. Brown A, Kirichek O, Balkwill A, et al. Comparison of dementia recorded in routinely collected hospital admission data in England with dementia recorded in primary care. *Emerg Themes Epidemiol*. 2016;13:11.

22. European Commission. *Revision of the European Standard Population. Report of Eurostat's Task Force*. Luxemburg; 2013. Available from: https://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f. Accessed February 7, 2019.

23. R Development Core Team. *R: A Language and Environment for Statistical Computing*; 2018. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/. Accessed February 7, 2019.

24. Williamson T, Miyagishima RC, Derochie JD, Drummond N. Manual review of electronic medical records as a reference standard for case definition development: a validation study. *CMAJ Open*. 2017;5(4):E830–E833.

25. Jaakkimainen RL, Bronskill SE, Tierney MC, et al. Identification of Physician-Diagnosed Alzheimer's disease and related dementias in population-based administrative data: a validation study using family physicians' electronic medical records. *J Alzheimers Dis*. 2016;54(1):337–349.

26. Herrett E, Shah AD, Boggon R, et al. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ*. 2013;346(3):f2350.

27. Bermejo-Pareja F, Benito-León J, Vega S, Medrano MJ, Román GC; Neurological Disorders in Central Spain (NEDICES) Study Group. Incidence and subtypes of dementia in three elderly populations of central Spain. *J Neurol Sci*. 2008;264(1–2):63–72.

28. Lobo A, Lopez-Anton R, Santabárbara J, et al. Incidence and lifetime risk of dementia and Alzheimer's disease in a southern European population. *Acta Psychiatr Scand*. 2011;124(5):372–383.

29. López-Pousa S, Vilalta-Franch J, Llinàs-Regla J, Garre-Olmo J, Román GC. Incidence of dementia in a rural community in Spain: the Girona cohort study. *Neuroepidemiology*. 2004;23(4):170–177.

30. Skinner T, Scott I, Martin J. Diagnostic errors in older patients: a systematic review of incidence and potential causes in seven prevalent diseases. *Int J Gen Med*. 2016;9:137–146.

31. Lang L, Clifford A, Wei L, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open*. 2017;7(2):e011146–011148.

32. Pi J, Olivé JM, Roca J, Masana L. Prevalence of dementia in a semi-rural population of Catalunya, Spain. *Neuroepidemiology*. 1996;15(1):33–41.

33. Gavrila D, Antúnez C, Tormo MJ, et al. Prevalence of dementia and cognitive impairment in southeastern Spain: the Ariadna study. *Acta Neurol Scand*. 2009;120(5):300–307.

34. García García FJ, Sánchez Ayala MI, A PM. [The prevalence of dementia and its main subtypes in subjects older than 65 years: impact of occupation and education. The Toledo Study]. *Med Clin (Barc)*. 2001;116(11):401–407. Spanish http://www.ncbi.nlm.nih.gov/pubmed/11333686.

35. Fernández M, Castro-Flores J, Perez-de Las Heras S, Mandaluniz-Lekumberri A, Gordejuela M, Zarranz J. [Prevalence of dementia in the elderly aged above 65 in a district in the Basque Country]. *Rev Neurol*. 2008;46(2):89–96. Spanish.

36. Lobo A, Launer LJ, Fratiglioni L, et al. Prevalence of dementia and major subtypes in Europe: a collaborative study of population-based cohorts. Neurologic diseases in the Elderly Research Group. *Neurology*. 2000;54(11 Suppl 5):S4–9.

37. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement*. 2013;9(1):63–75.

38. Bacigalupo I, Mayer F, Lacorte E. A systematic review and meta-analysis on the prevalence of dementia in Europe: estimates from the Highest-Quality studies adopting the DSM IV diagnostic criteria. *J Alzheimers Dis*. 2018;66(4):1471–1481.

39. Matthews FE, Arthur A, Barnes LE, et al. A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the cognitive function and Ageing Study I and II. *Lancet*. 2013;382(9902):1405–1412.

40. Gonçalves-Pereira M, Cardoso A, Verdelho A, et al. The prevalence of dementia in a Portuguese community sample: a 10/66 dementia research Group study. *BMC Geriatr*. 2017;17(1):261.

41. Kosmidis MH, Vlachos GS, Anastasiou CA, et al. Dementia prevalence in Greece: The Hellenic Longitudinal Investigation of Aging and Diet (HELIAD). *Alzheimer Dis Assoc Disord*. 2018;32(3):232–239.

42. Roehr S, Pabst A, Luck T, Riedel-Heller SG. Is dementia incidence declining in high-income countries? A systematic review and meta-analysis. *Clin Epidemiol*. 2018;10:1233–1247.

43. Langa KM, Larson EB, Crimmins EM, et al. A comparison of the prevalence of dementia in the United States in 2000 and 2012. *JAMA Intern Med*. 2017;177(1):51.

44. Lobo A, Saz P, Marcos G, et al. Prevalence of dementia in a southern European population in two different time periods: the ZARADEMP project. *Acta Psychiatr Scand*. 2007;116(4):299–307.

45. Craik FI, Bialystok E, Freedman M. Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve. *Neurology*. 2010;75(19):1726–1729.

46. Phung TK, Waltoft BL, Kessing LV, Mortensen PB, Waldemar G. Time trend in diagnosing dementia in secondary care. *Dement Geriatr Cogn Disord*. 2010;29(2):146–153.

47. Database - Eurostat. Available from: http://ec.europa.eu/eurostat/data/database. Accessed July 13, 2018.

48. Vilalta-Franch J, Garre-Olmo J, López-Pousa S, et al. [Feasibility of a telemedicine support system for diagnosing dementia in primary care]. *Rev Neurol*. 2012;55(5):263–269. Spanish.

49. Rait G, Walters K, Bottomley C, Petersen I, Iliffe S, Nazareth I. Survival of people with clinical diagnosis of dementia in primary care: cohort study. *BMJ*. 2010;341:c3584.

50. Sharp ES, Gatz M. Relationship between education and dementia: an updated systematic review. *Alzheimer Dis Assoc Disord*. 2011;25(4):289–304.

51. Gascón-Bayarri J, Reñé R, del Barrio JL, et al. Prevalence of dementia subtypes in El Prat de Llobregat, Catalonia, Spain: the PRATICON study. *Neuroepidemiology*. 2007;28(4):224–234.

Supplementary materials

Question: at this point, what is the basis for a diagnosis of dementia?

Answer 1- The diagnosis was made by a hospital specialist
- Indicate the subtype of dementia
- Indicate the severity of dementia

Answer 2- The diagnosis was based on cognitive and functional tests
- Specify the test(s) and the score(s)
- Indicate the subtype of dementia
- Indicate the severity of dementia
- Are the DSM-IV or CIE-10 diagnostic criteria fulfilled?
  (Yes/No)

Answer 3- The diagnosis was based on clinical opinion
- Symptoms were observed by a general practitioner, care giver or patient relative.
- Indicate the subtype of dementia
- Indicate the severity of dementia
- Are the DSM-IV or CIE-10 diagnosis criteria fulfilled?
  (Yes/No/Not sure)

Answer 4- The diagnosis is null (evolved or incorrect diagnosis)

Answer 5- Uncertain due to lack of information

Figure S1 Summary of questionnaire for the validation of dementia diagnosis.
Abbreviations: CIE-10, Código Internacional de Enfermedades (Spanish version of the ICD); DSM, *Diagnostic and Statistical Manual of Mental Disorders*.

Table S1 Characteristics of study population to estimate incidence rate of dementia

Characteristics	Patients with dementia, n (%), x [SD]	Dementia-free patients, n (%), x [SD]	Study population, n (%), x [SD]
Age	82.9 [7.0]	75.6 [7.9]	75.7 [7.9]
Women	6,218 (64.8)	575,442 (56.1)	581,660 (56.2)
Coronary heart disease	1,150 (12.0)	100,134 (9.8)	101,284 (9.8)
Cerebrovascular disease	1,854 (19.3)	92,722 (9.0)	94,576 (9.1)
Hypertension	6,972 (72.7)	653,792 (63.8)	660,764 (63.8)
Hyperlipidemia	5,076 (52.9)	549,597 (53.6)	554,673 (53.6)
Depression	2,676 (27.9)	163,728 (16.0)	166,404 (16.1)
Diabetes	2,868 (29.9)	251,599 (24.5)	254,467 (24.6)
Anemia	2,392 (24.9)	155,196 (15.1)	157,588 (15.2)
Hyperthyroidism	57 (0.6)	5,072 (0.5)	5,129 (0.5)
Cancer	1,996 (20.8)	193,459 (18.9)	195,455 (18.9)

Table S2 Crude incidence rates of dementia (per 1,000 person-years) in people aged ≥65 years in Catalonia (Spain) in 2016, stratified by sex and age

Sex	Age group (years)	Cases	Person/years	Incidence (95% CI)
Women	65–69	174	149,440	1.2 (1.0–1.4)
	70–74	522	129,528	4.0 (3.7–4.4)
	75–79	929	95,991	9.7 (9.1–10.3)
	80–84	1,739	93,356	18.6 (17.8–19.5)
	85–89	1,667	61,408	27.2 (25.9–28.5)
	≥90	1,187	38,113	31.1 (29.4–33.0)
Men	65–69	156	133,436	1.2 (1.0–1.4)
	70–74	398	111,178	3.6 (3.2–4.0)
	75–79	645	77,711	8.3 (7.7–9.0)
	80–84	928	65,940	14.1 (13.2–15.0)
	85–89	823	36,223	22.7 (21.2–24.3)
	≥90	428	16,430	26.1 (23.6–28.6)

Table S3 Characteristics of study population to estimate dementia prevalence that includes persons aged ≥65 years registered in SIDIAP on December 31, 2016

Characteristics	Patients with dementia, n (%), x [SD]	Dementia-free patients, n (%), x [SD]	Study population, n (%), x [SD]
Age	83.7 [6.9]	75.4 [7.7]	75.9 [7.9]
Women	43,995 (69.7)	555,947 (56.4)	599,942 (57.2)
Coronary heart disease	7,122 (11.3)	93,749 (9.5)	100,871 (9.6)
Cerebrovascular disease	12,281 (19.5)	85,878 (8.7)	98,159 (9.4)
Hypertension	44,880 (71.1)	628,783 (63.8)	673,663 (64.2)
Hyperlipidemia	34,324 (54.4)	533,427 (54.1)	567,751 (54.1)
Depression	18,389 (29.1)	157,376 (15.9)	175,765 (16.8)
Diabetes	18,361 (29.1)	239,853 (24.3)	258,214 (24.6)
Anemia	17,125 (27.1)	143,399 (14.6)	160,524 (15.3)
Hyperthyroidism	412 (0.7)	4,868 (0.5)	5,280 (0.5)
Cancer	11,914 (18.9)	179,050 (18.2)	190,964 (18.2)

Abbreviation: SIDIAP, System for the Development of Research in Primary Care.

Table S4 Crude prevalence of dementia in people aged ≥65 years in Catalonia (Spain) in 2016, stratified by sex and age

Sex	Age-group (years)	Cases	Population	Prevalence (95% CI)
Women	65–69	1,180	149,811	0.8 (0.7–0.8)
	70–74	2,767	131,252	2.1 (2.0–2.2)
	75–79	5,446	100,054	5.4 (5.3–5.6)
	80–84	11,493	102,434	11.2 (11.0–11.4)
	85–89	12,923	71,533	18.1 (17.8–18.3)
	≥90	10,186	44,858	22.7 (22.3–23.1)
Men	65–69	971	133,004	0.7 (0.7–0.8)
	70–74	2,045	111,636	1.8 (1.8–1.9)
	75–79	3,304	79,275	4.2 (4.0–4.3)
	80–84	5,413	69,060	7.8 (7.6–8.0)
	85–89	4,961	38,997	12.7 (12.4–13.1)
	≥90	2,428	17,042	14.2 (13.7–14.8)