Study Report

P4-C1-021

DARWIN EU® - Alzheimer's Disease: Incidence, Prevalence, and Individual's Characteristics

12/02/2026

Version 5.0

Authors: Rana Jajou, Marzyeh Amini, Melissa Leung, Talita Duarte-Salles, Adam Black, Ger Inberg, Maarten van Kessel, Ionna Nika, Cesar Barbosa, Ross Williams, Natasha Yefimenko

Public

CONTENTS

Study title	DARWIN EU® - Alzheimer's Disease: Incidence, Prevalence, and Individual's Characteristics
Study report version	V5.0
Date	12/02/2026
EUPAS number	EUPAS1000000826
Active substance	N/A
Medicinal product	N/A
Research question and objectives	The research questions are: • What are the overall incidence rates and prevalence of Alzheimer's disease in the general adult population, stratified by calendar year, sex, and age categories? • What are the demographic characteristics, diagnostic procedures, and clinical profile of individuals who are diagnosed with incident Alzheimer's disease? The specific objectives are: 1. To estimate overall incidence and prevalence of Alzheimer's disease in the general adult population, stratified by calendar year, sex, and age. 2. To describe the demographic characteristics, diagnostic procedures, and clinical profile of individuals who are diagnosed with incident Alzheimer's disease.
Countries of study	Croatia, Denmark, Germany, The Netherlands, The United Kingdom
Authors	Rana Jajou, r.jajou@darwin-eu.org Marzyeh Amini, m.amini@darwin-eu.org Melissa Leung, m.leung@darwin-eu.org Talita Duarte-Salles, t.duarte@darwin-eu.org

LIST OF ABBREVIATIONS

Acronyms/term	Description
Aβ	Amyloid beta
ACE	Angiotensin-converting enzyme
AD	Alzheimer’s disease
ADL	Activity of daily living
APOE4	Apolipoprotein E epsilon 4
APP	Amyloid Precursor Protein
ARB	Angiotensin II receptor blockers
ATC	Anatomical Therapeutic Chemical
ATHENA	OHDSI database for OMOP codes
CC	Coordination centre
CDM	Common Data Model
CHMP	Committee for Medicinal Products for Human Use (EMA)
CI	Confidence Interval
CIPH	Croatian Institute of Public Health
CPRD GOLD	Clinical Practice Research Datalink GOLD
DAC	Data Analytics Centre
DAR	Cause of Death Registry
DARWIN EU®	Data Analysis and Real-World Interrogation Network
DK-DHR	Danish Data Health Registries
DKMA	Danish Medicines Agency
DOAC	Direct oral anticoagulants
DOI	Declaration of Interests
DPP-4	Dipeptidyl Peptidase 4
DQD	Data Quality Dashboard
DRE	Digital Research Environment
EBM	Einheitlicher Bewertungsmaßstab
EHR	Electronic Health Records
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
EUPAS	EU Post-Authorisation Studies Register
GmbH	Gesellschaft mit beschränkter Haftung
GDPR	General Data Protection Regulation
GLP-1	Glucagon-like peptide-1
GP	General Practitioner

Acronyms/term	Description
HMA	Heads of Medicines Agencies
HRI	Health Risk Institute
ICD-10	International Classification of Diseases, Tenth Revision
ICPC-1	International Classification of Primary Care, first version
ICU	Intensive Care Unit
IgG1	Immunoglobulin G1
InGef RDB	InGef Research Database (Germany)
IP	Inpatient
IPCI	Integrated Primary Care Information (the Netherlands)
IQR	Interquartile Range
IQVIA DA	IQVIA Disease Analyzer
IRB	Institutional Review Board
LOINC	Logical Observation, Identifiers, Names and Codes
MCI	Mild Cognitive Impairment
MRI	Magnetic Resonance Imaging
NA	Not applicable
NAJS	National Public Health Information System (Croatia)
NDORMS	Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences
NFC	Near Field Communication
NHG	Nederland Huisartsen Genootschap
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
ONS	Office for National Statistics
OP	Outpatient
OPS	Operationen- und Prozedurenschlüssel
PCSK9	Proprotein convertase subtilisin/kexin type 9
PET-F18	Positron Emission Tomography with Fluorine-18
PSEN1	Presenilin-1
PSEN2	Presenilin-2
PY	Person-years
PZN	Pharmazentralzummer -pharmaceutical reference number
RWD	Real-World Data
RWE	Real-World Evidence
RxNorm	Medical prescription normalized
SD	Standard Deviation
SGLT2	Sodium-glucose cotransporter-2
SHI	Statutory Health Insurances

Acronyms/term	Description
SNOMED	Systematized Nomenclature of Medicine
UK	The United Kingdom
WHO	World Health Organisation

1. TITLE

DARWIN EU® - Alzheimer's Disease: Incidence, Prevalence, and Individual's Characteristics

2. DESCRIPTION OF THE STUDY TEAM

Study team role	Names	Organisation
Principal Investigators	Rana Jajou Marzyeh Amini Melissa Leung	Erasmus MC
Co-Principal Investigator	Talita Duarte-Salles	Erasmus MC
Data Scientists	Adam Black Ger Inberg Maarten van Kessel Ionna Nika Cesar Barbosa Ross Williams	Erasmus MC
Study Manager	Natasha Yefimenko	Erasmus MC
Data source	**Names**	**Data Partner Organisation***
NAJS	Antea Jezidžić Marko Čavlina Karlo Pintaric Anamaria Jurcevic Jakov Vukovic	Croatian Institute of Public Health
DK-DHR	Elvira Bräuner Susanne Bruun	Danish Medicines Agency
InGef RDB	Josephine Jacob Raeleesha Norris Alexander Harms Annika Vivirito	Institut für angewandte Gesundheitsforschung Berlin GmbH
IQVIA DA Germany	Dina Vojinovic Akram Sharim Mendez Rangel Ellen Gerritsen Gargi Jadhav Hugo Vernooij Isabella Kaczmarczyk	IQVIA

IPCI	Katia Verhamme	Erasmus MC
CPRD GOLD	Antonella Delmestri	University of Oxford

*Data partners do not have an investigator role. Data partners execute code at their data source, review, and approve their results.

3. ABSTRACT

Title

DARWIN EU® - Alzheimer's Disease: Incidence, Prevalence, and Individual's Characteristics

Rationale and background

Alzheimer's disease (AD) is the most common cause of dementia, characterised by progressive cognitive decline and loss of independence. Its burden is rising globally due to population ageing, with substantial implications for healthcare systems and society. Recent therapeutic advances, such as the approval of anti-amyloid monoclonal antibodies, are reshaping AD management but also highlight the need for comprehensive and updated real-world evidence on the epidemiology and characteristics of affected populations.

This DARWIN EU® study provides contextual information through an overview of the incidence and prevalence of AD in the general adult population in European countries, as well as demographic and clinical characteristics of individuals with AD.

Research question and objectives

Research questions

What is the overall incidence rate and prevalence of AD? What are the demographic characteristics, diagnostic procedures, and clinical profiles of individuals who are diagnosed with incident AD?

Objectives

1. To estimate overall incidence and prevalence of AD in the general adult population, stratified by calendar year, sex, and age.
2. To describe the demographic characteristics, diagnostic procedures, and clinical profile of individuals who are diagnosed with incident AD.

Methods

Study design

Cohort descriptive study.

Population

Objective 1: All individuals aged 18 years or older present in the data sources during the study period from 01/01/2014 to 31/12/2024, or until earliest date of loss to follow-up, death, end of the observation period or data availability, or until the earliest date of AD diagnosis with at least 365 days of database history.

Objective 2: All individuals aged 18 years or older, who are diagnosed with incident AD during the study period from 01/01/2014 to 31/12/2024, with at least 365 days of database history.

Variables

Exposure:

Not applicable.

Outcome:

Alzheimer's disease.

Relevant covariates:

At the diagnosis date, covariates included sex, age categories (18–65, ≥66, 18–55; 56–65; 66–75; 76–85; ≥86), and calendar year.

Within 365 days prior to AD diagnosis, covariates included diagnostic procedures, such as recording of brain MRI and recording of brain PET-F18, predefined clinical features, including alterations in activities of daily living (ADLs), and caregiver support, as well as medication use. The latter included AD drugs and medications for comorbidities, such as antiplatelets, anticoagulants/antithrombotic, glucose-lowering therapies (insulin and oral agents), antihypertensives, antiarrhythmics/rhythm control drugs, lipid-lowering drugs.

At any time prior and up to AD diagnosis, covariates included predefined comorbidities of interest, such as Down's syndrome, stroke, atrial fibrillation, myocardial infarction, heart failure, hypertension, diabetes, hypercholesterolemia, and hypertriglyceridemia. In addition, the occurrence of a prior diagnosis of Mild Cognitive Impairment (MCI) was described.

Data sources

1. Croatia: Croatian National Public Health Information System (NAJS)
2. Denmark: Danish Data Health Registries (DK-DHR)
3. Germany: InGef Research Database (InGef RDB)
4. Germany: IQVIA Disease Analyzer Germany (IQVIA DA Germany)
5. The Netherlands: Integrated Primary Care Information (IPCI)
6. The United Kingdom, Clinical Practice Research Datalink GOLD (CPRD GOLD)

Study size

No sample size was calculated, as this is a descriptive study which did not test a specific hypothesis. Based on a preliminary feasibility assessment, the expected number of persons counts for AD in the data sources included in this study ranged from 13,400 (IPCI) to 109,700 (DK-DHR).

Statistical analysis

Objective 1: Incidence rates per 100,000 person-years (PY) and prevalence of AD were estimated, overall and stratified by calendar year, sex, and age categories. Incidence rates and prevalence were reported with 95% confidence intervals.

Objective 2: Demographic characteristics, pre-specified comorbidities, concomitant medications, diagnostic procedures, and prior cognitive diagnosis were described and reported as counts and proportions. For MCI, time from MCI recording to AD diagnosis was described.

A minimum cell counts of 5 was used when reporting results, with any smaller count reported as "<5" and zero counts as "0".

Results

Across six European data sources, individuals with AD were consistently older (median age at index date 79–82 years) and predominantly female (56.2–66.4%) compared to the general adult population (median age 34–51 years, 49.4–55.9% female).

The overall incidence rates of AD per 100,000 PYs in the general adult population (≥18 years old) ranged from 53.85 in IQVIA DA Germany to 137.3 in InGef RDB over the whole study period. In general, the incidence was higher among females than males and increased with age across all data sources. The overall incidence rate, over the whole study period, of early-onset AD (18–65 years of age, ranging from 3.71 to

8.72) was lower than late-onset AD (≥66 years of age, ranging from 161.14 to 742.40) across the data sources.

The overall prevalence of AD ranged from 0.2% in IQVIA DA Germany to 1.7% in InGef RDB over the whole study period. The prevalence of AD was higher among females than males and increased with age across all data sources. Over the whole study period, the overall prevalence of early-onset AD was lower than 0.10% across all data sources, and for late-onset AD, the prevalence ranged from 0.5–5.1%.

The prevalence of comorbidities (at any time prior and up to AD diagnosis) differed widely across data sources. The comorbidities that were highly prevalent among AD cases were hypertension (30.9–83.9%) and diabetes (13.1–29.7%). Cardiovascular diseases were also common, with atrial fibrillation ranging from 4.1–20.3% and heart failure ranging from 3.7–21.9%. Use of diagnostic procedures (in the year prior AD diagnosis) varied markedly, with brain MRI ranging from 13.9–18.0% and PET-F18 scans rarely recorded except in DK-DHR (6.6%). Clinical characteristics (in the year prior AD diagnosis) also differed between data sources, with caregiver support reported in 0.7–21.1% of cases, daily activity limitations reported in <5 cases–7.8%, and prior MCI diagnoses recorded in 1.9–4.4% cases. AD medication use varied notably between data sources, with memantine (0.4–31.2%) and donepezil (7.1–23.7%) showing the greatest variation, while rivastigmine and galantamine use remained consistently low (0.0–6.0%). Treatments for comorbidities also varied between data sources, including antiplatelets (1.4–38.3%), anticoagulants (5.6–17.4%), antihypertensives (>40.0% in most sources), and heart-failure therapies (35.2–77.0%), which were particularly frequent in NAJS (68.4%) and InGef RDB (77.0%). Among individuals with prior MCI, time to AD diagnosis varied considerably, with median intervals from approximately 350 to over 600 days, longer in females and increasing with age.

Discussion

Overall, the findings highlight consistent patterns with respect to the incidence and prevalence estimates by age (higher estimates among older age) and sex (higher estimates among females). Furthermore, consistent demographic patterns of individuals with AD were observed across the data sources. The differences in ranges of incidence and prevalence estimates between data sources may be partially explained by the differences in data source types, i.e., inpatient versus outpatient data, and differences in the mapping of the source vocabulary to the Observational Medical Outcomes Partnership Common Data Model.

AD populations were in general older (median 79–82 years) and more frequently female (56.2–66.4%) than the general adult population. Among the AD population, hypertension and diabetes were highly prevalent. Diagnostic procedures were rarely used or not available in the data sources. This was also the case for the prespecified clinical features. Therefore, these variables might not be truly representative and should be interpreted with caution. The most common used AD drugs were memantine and donepezil. The median time from MCI to AD diagnosis ranged from approximately 350 to over 600 days and increased with increasing age.

The data sources included were selected based on several reasons, including the possibility to operate under blanket Institutional Review Board (IRB) approval or that approval could be obtained within a month to ensure the feasibility of the study execution within the study timelines. Future studies can include a wider range of data sources (which require longer IRB approval) and countries. This will allow broader generalisability of the findings and, possibly, the inclusion of data sources with more detailed information on demographic characteristics, diagnostic procedures, and clinical profile of the AD population.

4. AMENDMENTS AND UPDATES

None.

5. MILESTONES

Study deliverable	Timelines (planned)	Timelines (actual)
Final Study Protocol	12 November 2025	12 November 2025
Creation of Analytical code	24 October 2025	24 October 2025
Execution of Analytical Code on the data	6 November 2025	6 November 2025
Draft Study Report	24 November 2025	8 December 2025
Final Study Report	To be confirmed by EMA	To be confirmed by EMA

6. RATIONALE AND BACKGROUND

Alzheimer's disease (AD) is a disorder that causes degeneration of the cells in the brain, and it is the main cause of dementia, which is characterised by a decline in thinking and independence in personal daily activities.[1] AD is considered a multifactorial disease associated with several risk factors, such as increasing age, genetic factors, head injuries, vascular diseases, infections, and environmental factors.[1]

The burden of AD is increasing globally, with significant implications for public health and the economy.[2-4] Demographic analyses suggest that these patterns are driven by decreases in fertility coupled with increases in life expectancy, which together lead to large changes in the age structure of the population (i.e., larger number of people at the oldest ages). These changes have led to increases in the number of people affected by dementia, including AD, over time.[2, 4] In 2021, global mortality from AD and other dementias among individuals aged 60 years and older reached approximately 1,922,970 cases (95% confidence interval [CI]: 480,348 to 5,104,315), and the prevalence was 52,560,253 cases (95% CI: 41,399,948 to 65,633,448). Projections suggest a near fourfold increase in AD cases by 2050, driven by population growth and aging, with females disproportionately affected.[5]

This DARWIN EU® study was proposed to provide contextual information through an overview of the prevalence and incidence of AD, as well as individual's demographics and clinical characteristics in European countries in the last decade.

7. RESEARCH QUESTION AND OBJECTIVES

Research question

What is the overall incidence rate and prevalence of AD in the general adult population, stratified by calendar year, sex, and age categories?

What are the demographic characteristics, diagnostic procedures, and clinical profiles of individuals who are diagnosed with incident AD?

Objectives

This study aimed to provide contextual information on the incidence and prevalence of AD in the general population and demographic- and clinical characteristics of individuals with newly diagnosed AD.

The specific objectives of this study were:

1. To estimate overall incidence and prevalence of AD in the general adult population, stratified by calendar year, sex, and age.

2. To describe the demographic characteristics, diagnostic procedures, and clinical profile of individuals who are diagnosed with incident AD.

8. RESEARCH METHOD

8.1. Study design

Population level and patient level cohort studies were conducted using routinely collected health data from six data sources across five European countries.

- A characterisation study including the general population was conducted to address objective 1.
- A characterisation study including individuals diagnosed with incident AD was conducted to address objective 2.

A graphical description of the study design for each objective is shown in **Figure 1** and **Figure 2**.

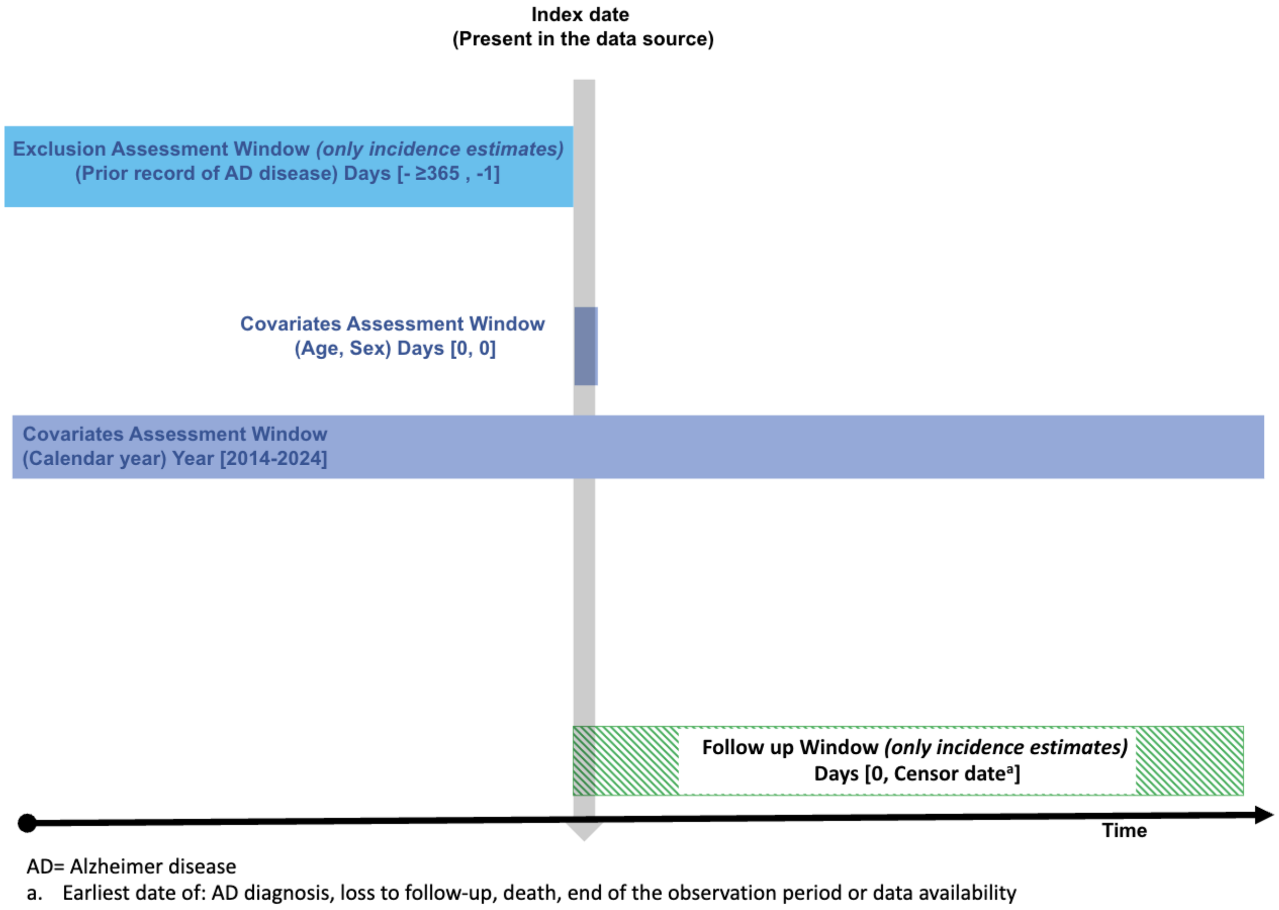

Figure 1. Graphical depiction of the study design (Objective 1).

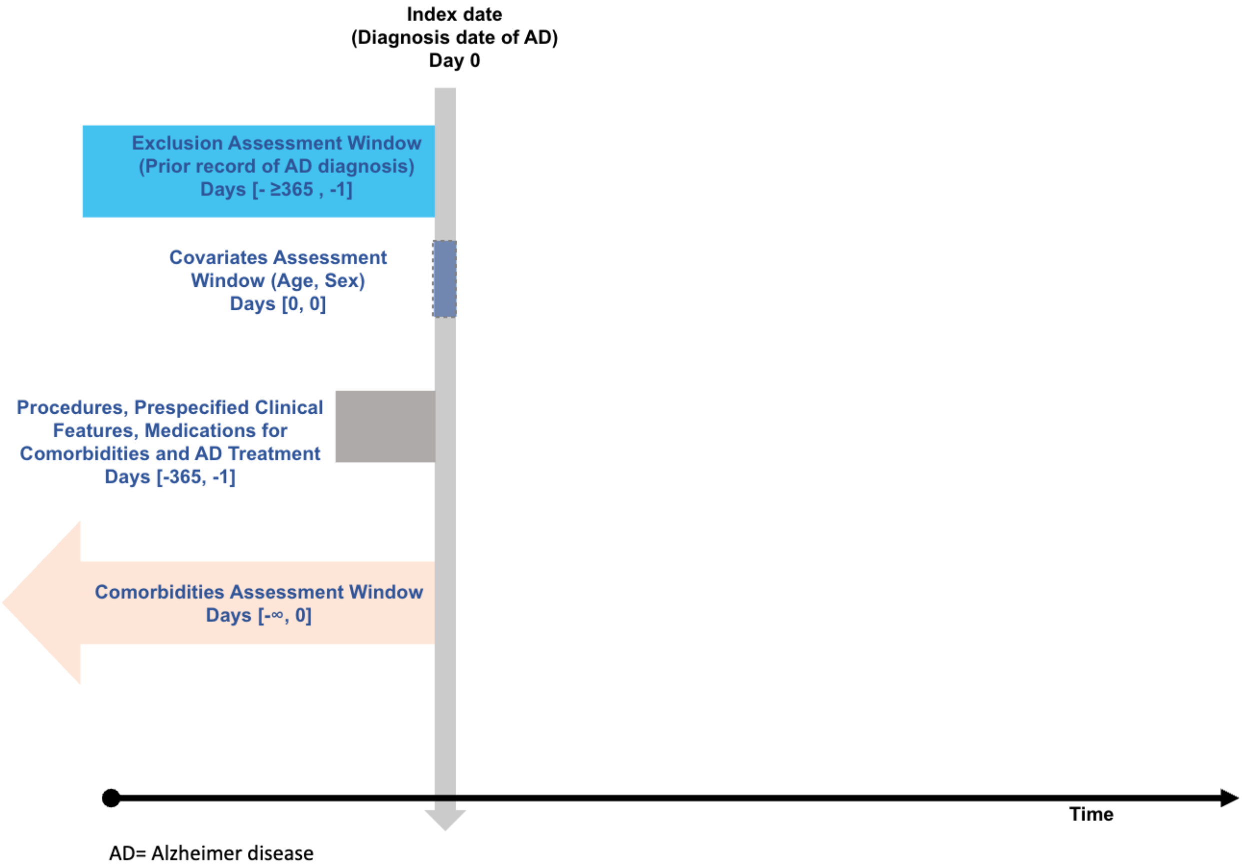

Figure 2. Graphical depiction of the patient-level study design (Objective 2).

8.2. Follow-up

The follow-up was defined as follows for each cohort:

For the cohort of the general adult population (Objective 1):

The index date was the first date during the study period, on which individuals were minimum 18 years of age and had sufficient prior data availability (minimum 365 days).

Follow-up started on the study start date (01/01/2014) or the date on which the requirement of prior history was fulfilled.

Follow-up ended on the earliest date of AD diagnosis, loss to follow-up, death, end of observation period (the latest available data, due to migration, deregistration at general practitioner [GP] practice or health insurance company, but also death and loss to follow-up), or study end date of 31/12/2024.

For the cohort of individuals who are diagnosed with incident AD (Objective 2):

The index date was the date of first (incident) AD diagnosis, on which individuals were minimum 18 years of age and had sufficient prior data availability (minimum 365 days). From the index date, retrospective data was assessed on comorbidities, diagnostic procedures, prespecified clinical features, and medications for comorbidities and AD treatment.

An example of entry and exit into the denominator population is shown **Figure 3**. In this example, person ID 1 already had sufficient prior history before the start date and the observation period ended after the study end date, so this person contributed during the complete study period. Person IDs 2 and 4 entered the study only when they had sufficient prior history. Person ID 3 exited the data source on the end of the observation period. Lastly, person ID 5 had two observation periods in the data source. In the first period, the person contributed time from the study start until the end of the observation period. In the second period, the person started contributing time again once sufficient prior history was reached and exited at study end date.

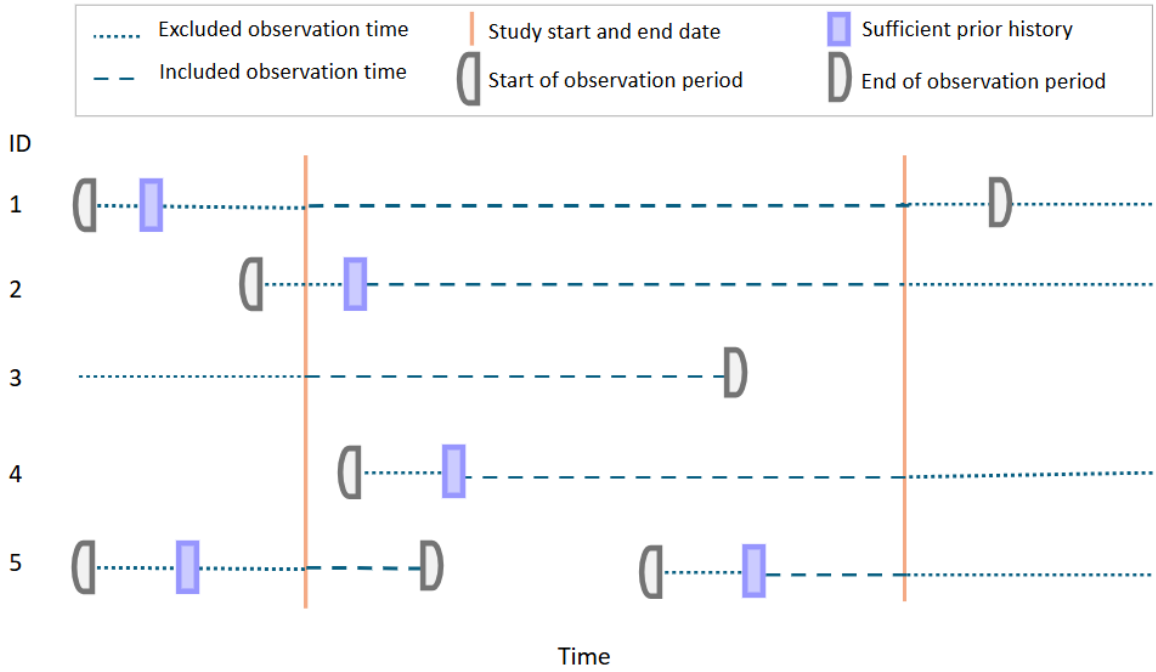

Figure 3. Included observation time for the denominator population.

8.3. Study population with inclusion and exclusion criteria

The study population was defined as follows for each cohort:

For the cohort of the general adult population (Objective 1):

Inclusion criteria

- All individuals present during the study period
- Minimum of 365 days of available data history
- Individuals were aged ≥18 at index date

Exclusion criteria

- For incidence estimates: individuals with any prior record of AD diagnosis before index date.

For the cohort of individuals who are diagnosed with incident AD (Objective 2):

Inclusion criteria

- Individuals who are diagnosed with incident AD present during the study period
- Minimum of 365 days of continuous prior observation before the AD index date
- Individuals were aged ≥18 at index date

Exclusion criteria

- Individuals with any prior record of AD diagnosis before index date.
- Individuals with unspecified dementia.

8.4. Study setting and data sources

This study was conducted using routinely collected data from six primary/secondary care data sources in the DARWIN EU® network of data partners from five European countries. All data were a priori mapped to the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM). Information on data sources included and a rationale for their choice in terms of ability to capture the relevant data is described in **Table 1** and **ANNEX I and II**.

Table 1. Description of the selected data sources.

Country	Name of Data source	Health Care setting	Type of Data	Number of active individuals	Calendar period covered by each data source	Contributing to
Croatia	National Public Health Information System (NAJS)	Primary care: GPs, secondary care: specialists (ambulatory or hospital outpatient care), hospital inpatient care	Registry	4.3M	2017–2024	Objectives 1 and 2
Denmark	Danish Data Health Registries (DK-DHR)	Community pharmacists, secondary care: specialists (ambulatory or hospital outpatient care), hospital inpatient care	Registry	5.98M	1995–2024	Objectives 1 and 2
Germany	InGef Research Database (InGef RDB)	Primary care: GPs, community pharmacists, primary care specialists (e.g., paediatricians), secondary care: specialists (ambulatory or hospital outpatient care), hospital inpatient care, claims data	Claims	7.67M	2015–2024	Objectives 1 and 2
Germany	IQVIA Disease Analyzer Germany (IQVIA DA Germany)	Primary care: GPs, primary care specialists (e.g., paediatricians)	Outpatient General Practitioner Care	4.48M	1992–2024	Objectives 1 and 2

Country	Name of Data source	Health Care setting	Type of Data	Number of active individuals	Calendar period covered by each data source	Contributing to
The Netherlands	Integrated Primary Care Information (IPCI)	Primary care: GPs	Outpatient General Practitioner Care	1.33M	2006–2024	Objectives 1 and 2
The United Kingdom	Clinical Practice Research Datalink (CPRD) GOLD	Primary care	EHR	2.83M	1988–2024	Objectives 1 and 2

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
GP=General Practitioner, M=millions, EHR=Electronic Health Records

Data sources

1. Croatia: Croatian National Public Health Information System (NAJS)
2. Denmark: Danish Data Health Registries (DK-DHR)
3. Germany: InGef Research Database (InGef RDB)
4. Germany: IQVIA Disease Analyzer Germany (IQVIA DA Germany)
5. The Netherlands: Integrated Primary Care Information (IPCI)
6. The United Kingdom, Clinical Practice Research Datalink GOLD (CPRD GOLD)

Data sources selection

These data sources fulfil the criteria required in terms of data quality, completeness, timeliness, and representativeness for this study while covering different regions of Europe (**ANNEX I and II**). Furthermore, the data sources selected operate under blanket IRB approval or can obtain approval within a month, ensuring feasibility of study execution within the current timelines.

Data source justification and key characteristics

Croatia, National Public Health Information System (NAJS)

NAJS includes primary care, outpatient specialist care, and inpatient care registries providing information on AD occurrence, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18 years). The CDM population comprises all publicly insured (basic insurance is mandatory for all citizens and residents in Croatia) persons residing in Croatia starting in 2017. The inclusion of NAJS enhances the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimated approximately 28,900 individuals with AD in NAJS.

Data availability and follow-up were sufficient, with records available from 01/01/1995 (accurate data are available starting from the year 2017) and the most recent data extraction on 02/08/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, IRB approval for NAJS was obtained within approximately one month, supporting the feasibility of study execution within the study timelines.

Denmark, Danish Data Health Registries (DK-DHR)

DK-DHR includes nationwide registry data, comprising inpatient, outpatient, and emergency care data from hospitals, providing information on AD occurrence, radiology, medications, and relevant comorbidities in the general adult population (≥18 years). However, caregiver support and activities of daily living were not available. The inclusion of DK-DHR enhances the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimates approximately 107,900 individuals with AD in DK-DHR. The absence of primary care data meant that behavioural and neuropsychiatric symptoms were not detected during this study if they were treated only in primary care. Moreover, no data was available on cognitive test results and functional status. Brain MRI was included in MRI of head, but it was not mapped on a more granular level.

Data availability and follow-up were sufficient, with records available from 01/01/1995 and the most recent data extraction on 10/04/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, DK-DHR operates under blanket IRB approval, ensuring feasibility of study execution within the current timelines.

Germany, InGef Research Database (InGef RDB)

InGef RDB includes primary care, hospital inpatient care, and secondary outpatient care claims data, providing information on AD occurrence, radiology, medications, and relevant comorbidities in the general adult population (≥18 years). However, activities of daily living, caregiver support, and review of mild cognitive impairment were not available. The CDM population comprises individuals from all regions in Germany who are covered by one of the approximately 50 contributing statutory health insurance (SHI) companies, ~15% of the German population starting in 2015. Approximately 90% of the German population is insured through SHI. Those not publicly insured include individuals with specific exemptions from the general requirement for enrolment in public insurance. This includes individuals who earn a high income and can opt-out of the system to seek private insurance, freelancers/self-employed people who prefer to seek private insurance, civil servants (teachers, police, and other government employees), some foreign nationals, etc. While InGef RDB includes individuals across the SES spectrum and while the vast majority of the population is publicly insured, the fact that privately insured groups are not included might introduce a certain degree of selection bias. The inclusion of InGef RDB enhanced the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimated approximately 27,500 individuals with AD in InGef RDB.

Data availability and follow-up were sufficient, with records available from 01/01/2015 and the most recent data extraction on 18/04/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, InGef RDB operates under blanket IRB approval, ensuring feasibility of study execution within the current timelines.

Germany, IQVIA Disease Analyser (IQVIA DA Germany)

IQVIA DA Germany includes primary and outpatient secondary care electronic health record data, providing information on AD occurrence, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18 years). IQVIA DA Germany has a nationwide scope via a representative EHR panel of outpatient GP and office-based specialist practices sampled across all federal states. Patients are captured regardless of insurer (SHI/GKV and private/PKV), with insurer fields available for stratification where recorded. The inclusion of IQVIA DA Germany enhanced the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimated approximately 79,300 individuals with AD in IQVIA DA Germany.

Data availability and follow-up were sufficient, with records available from 01/01/1992 and the most recent data extraction on 10/04/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, IQVIA DA Germany operates under blanket IRB approval, ensuring feasibility of study execution within the current timelines.

The Netherlands, The Integrated Primary Care Information (IPCI)

IPCI includes primary care electronic health records data, providing information on AD occurrence, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18 years). The inclusion of IPCI enhanced the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimated approximately 13,400 individuals with AD in IPCI.

Data availability and follow-up were sufficient, with records available from 01/01/2006 and the most recent data extraction on 16/04/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, IRB approval for IPCI was obtained within approximately one month, ensuring feasibility of study execution within the current timelines.

Clinical Practice Research Datalink GOLD, the United Kingdom (CPRD GOLD)

CPRD GOLD includes primary care-GP, primary specialist care, secondary specialist care of outpatient electronic health records, providing information on AD occurrence, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18 years). The CDM population comprises all persons residing in the United Kingdom starting in 2015. The inclusion of CPRD GOLD enhanced the geographical diversity of data sources, with adequate data coverage over the study period. A preliminary feasibility assessment estimated approximately 87,500 individuals with AD in CPRD GOLD.

Data availability and follow-up were sufficient, with records available from 01/01/1988 and the most recent data extraction on 15/03/2025, fully aligned with the study period. No study-specific limitations were identified for this data source. Furthermore, CPRD GOLD operates under blanket IRB approval, ensuring feasibility of study execution within the current timelines.

8.5. Study period

The study period was from 01/01/2014 to 31/12/2024 or the most recent data available for each contributing data source. It should be noted that in the NAJS data source, accurate data were available starting from 01/01/2017 (see **ANNEX II**), while in InGef RDB, data were available starting from 01/01/2015. Therefore, to meet the 365 days prior history requirement, the study period for NAJS was from 01/01/2018 to 31/12/2024; and for InGef RDB from 01/01/2016 to 31/12/2024.

8.6. Variables

8.6.1. Exposure

None.

8.6.2. Outcome

For objective 1, the outcome was defined as the occurrence of AD recorded in the data sources.

The definition of Alzheimer’s disease was based on SNOMED codes. Diagnostic codes explicitly referring to AD or primary degenerative dementia of the Alzheimer type (including presenile/senile onset, uncomplicated, or with depression, delirium, delusions, or behavioural disturbance) were included. The concept sets used for the identification of individuals with AD are described in **ANNEX III.** These codes were refined during the study execution following the DARWIN EU® phenotyping standard processes,[6] which involved the review of code lists by clinical experts and the review of phenotypes after their execution in the participating data sources.

For objective 2, AD occurrence was part of the inclusion criteria. There was no outcome for objective 2.

8.6.3. Covariates, including confounders, effect modifiers, intercurrent events, and other variables

All objectives:

The following covariates were considered:

- Sex
    - Female/male
- Age at index date was calculated using January 1st of the year of birth as proxy for the actual date of birth. Age groups were categorized categorised considering that early-onset AD is usually identified in individuals aged 65 years and younger and late-onset AD in individuals aged older than 65 years, namely:
    - 18–55 years
    - 56–65 years
    - 66–75 years
    - 76–85 years
    - ≥86 years

Objective 1:

- Calendar year: Calendar time was based on the calendar year of AD incidence and was from 2014 to 2024.
- Early onset AD (18–65 years) versus late-onset AD (≥66 years).

Objective 2:

The diagnostic procedures and clinical profile of individuals with newly diagnosed AD were assessed in different time windows of interest as follows:

- Within 365 days prior to diagnosis:
    - Recording of diagnostic procedures:
        - Brain MRI (date of test)
        - Brain PET-F18 (date of test)

      For these procedures, only information on whether the test was recorded/performed (yes/no) derived from the date of the test were described. Of note, information on test results was not consistently captured across data sources.
    - Clinical features:
        - Alterations in activities of daily living (ADLs): any changes, limitations, or dependencies related to an individual’s ability to perform basic or instrumental self-care tasks. This category encompasses findings such as alteration in ADL, need for assistance with ADLs, difficulty performing personal hygiene activities, functional dependency, physical functional dependency, unfitness for activity, and limitations in instrumental activities of daily living.
        - Caregiver support: based on recordings of clinical findings related to caregiver support: problems with life management, an individual’s dependence on a care provider, or the need for personal care assistance.

    - Alzheimer's disease drugs: include memantine, donepezil, rivastigmine, galantamine. The code list of products is shown in **ANNEX III**.
    - Medications for predefined comorbidities including antiplatelets, anticoagulants/antithrombotic, glucose-lowering therapies (insulin and oral agents), antihypertensives, antiarrhythmics/rhythm control drugs, lipid-lowering drugs.

  If, during study execution, the counts for ADLs were found to be too low to yield reliable estimates, these variables were excluded from the final analysis and reported as deviation for the protocol with the corresponding justification.

- At any time prior and up to diagnosis:
  - Predefined comorbidities of interest:
    - Down's syndrome
    - Stroke
    - Atrial fibrillation
    - Myocardial infarction
    - Heart failure
    - Hypertension
    - Diabetes
    - Hypercholesterolemia
    - Hypertriglyceridemia
  - Prior cognitive diagnosis:
    - Mild Cognitive Impairment (MCI) recorded before AD diagnosis

The concept sets used for the identification of covariates are described in **ANNEX III**. All codes were refined during the study execution following the DARWIN EU® phenotyping standard processes,[6] which involved the review of code lists by clinical experts and the review of phenotypes after their execution in the participating data sources.

8.7. Study size

No sample size was calculated, as this was a descriptive disease epidemiology study which did not test a specific hypothesis. In addition, the study was based on secondary use of data (i.e., data already collected for other purposes than research) to estimate incidence and prevalence of AD in the general adult population. Thus, the sample size was driven by the availability of data of individuals with AD. Based on a preliminary feasibility assessment, the expected number of person counts for AD diagnosis in the data sources included in this study ranged from 13,400 (IPCI) to 109,700 (DK-DHR). These numbers were based on the overall number of conditions or observations registered in each data source with no filter by study period or inclusion and exclusion criteria.

8.8. Data transformation

Analyses were conducted separately for each data source. Before study initiation, test runs of the analyses were performed on a subset of the data sources and quality control checks were performed. Once all the tests were passed (see **ANNEX IV**), the final study codes package was released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally executed the analytics against the OMOP CDM in R Studio and reviewed and approved the, by default, aggregated results.

The study results of all data sources were checked, after which they were made available to the team, and the dissemination phase started. All results were locked and timestamped for reproducibility and transparency.

8.9. Statistical methods

8.9.1. Main summary measures

The R packages *IncidencePrevalence*,[7] *CohortCharacteristics*,[8] and *PatientProfiles*,[9] developed by DARWIN EU®, were used for the population level characterisation, individual-level characterisation, and large-scale characterisation, based on OMOP CDM mapped data.

The main measures used in the analyses included frequencies (counts and percentages) for categorical variables; mean, standard deviation (SD), median, interquartile range (IQR), and range for continuous variables; incidence rates (IR) per 100,000 person-years (PYs) with 95% Poisson confidence intervals (CIs) to assess the newly diagnosed AD individuals; and prevalence with 95% Binomial CIs to assess the number of individuals with a recorded AD diagnosis.

8.9.2. Main statistical methods

Objective 1 (Incidence and prevalence of AD)

Incidence calculations

Yearly incidence rates of AD were calculated as the number of newly diagnosed cases of AD divided by the total person-time at risk, expressed per 100,000 PYs, for each calendar year. Eligible participants contributed person-time from cohort entry until the earliest of the following: first AD diagnosis during the study period, death, end of observation, or end of the study period (31/12/2024). Participants without a diagnosis contributed time at risk until censoring, as described in section **8.2 Follow-up**. The 95% confidence intervals were based on a Poisson distribution.

An illustration of the calculation of incidence of AD is shown above in **Figure 4**. Person IDs 1 and 4 contributed person-time at risk from the study start date until the occurrence of their first recorded AD diagnosis (incident case). Person ID 3 was excluded from the analysis due to a previous record of AD disease. Person IDs 2 and 5 contributed person-time at risk from the study start date until the study end date, as no AD diagnosis was observed during the study period or before study entry. Periods of observation prior to eligibility were excluded from the time at risk.

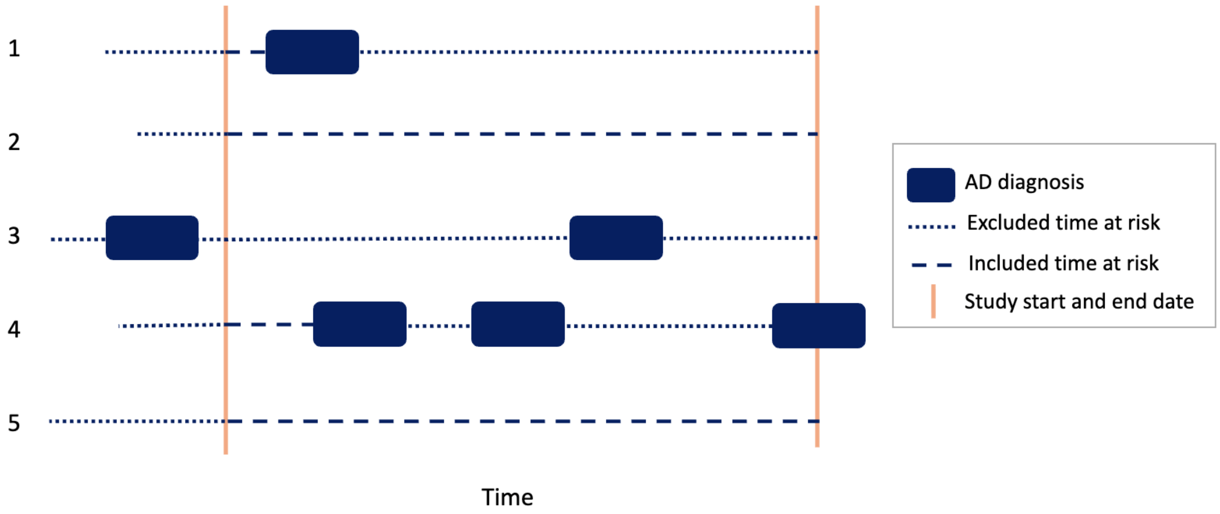

Figure 4. Examples of incidence for Alzheimer's Disease (AD).

Prevalence calculations

Prevalence was calculated as annual period prevalence which summarises the number of individuals with a recorded diagnosis of AD who were alive and actively observed at a given calendar year. Therefore, period prevalence gives the proportion of individuals exposed at any time during a specified interval. Binomial 95% confidence intervals were calculated.

An illustration of the calculation of period prevalence is shown below in **Figure 5**. Between time *t+2* and *t+3*, all five participants contributed observation time, and two of the five study participants had a recorded AD diagnosis, giving a prevalence of 40%. For the period *t* to *t+1*, all five participants contributed observation time, with one of the five having an AD diagnosis, giving a prevalence of 20%.

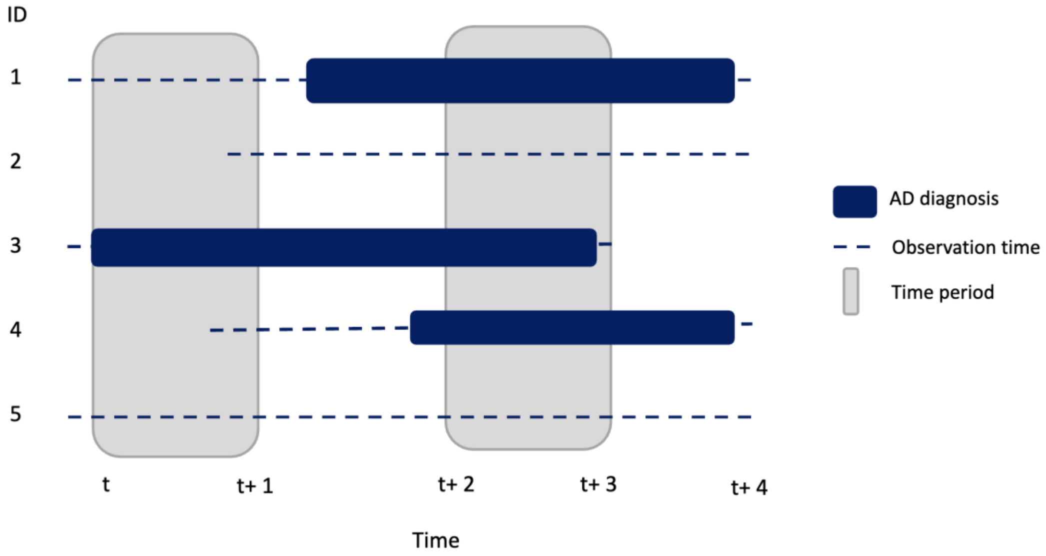

Figure 5. Example of Period Prevalence Calculation for Alzheimer's Disease (AD).

Objective 2 (Diagnostic procedures, individual-level, and clinical profile characterisation)

Descriptive statistics were used to summarise demographic and clinical characteristics at different time windows as described in **Section 8.6.3.**

Categorical variables (e.g. sex, comorbidities, medication use) were described using counts and percentages.

Continuous variables (e.g. age at diagnosis) were described using means, standard deviations, medians, interquartile ranges, and ranges (min–max).

For individuals with a recorded diagnosis of MCI prior to their first AD diagnosis, the absolute/relative frequency, as well as the time elapsed between MCI occurrence and AD diagnosis, was calculated. The duration was defined as the time in days between the date of the earliest recorded MCI diagnosis and the index date of AD diagnosis. If multiple MCI records were present, the first available record was used. Time elapsed between recording of MCI and AD diagnoses was summarised as descriptive statistics (median, IQR, mean, standard deviation) and presented both overall and stratified by sex and age categories.

8.9.3. Missing values

Missing data were not imputed.

8.9.4. Sensitivity analysis

No sensitivity analyses were performed.

8.10. Deviations from the protocol

Deviation number	Protocol version	Date	Section of study protocol	Deviation	Reason
1	V3.0 (Final)	12 November 2025	**Section 8.6.3.**	Information on Neuropsychiatric symptoms (as part of the 'Prespecified Clinical Features') is not provided.	This data is not captured in the OMOP version of the data sources. Further explorative work would be needed in the original data sources to determine the availability of this information and whether it is captured in structured or unstructured (e.g., free text) format.

9. RESULTS

The full set of results for this study is available through an interactive web-application ShinyApp at EUPAS1000000826.

9.1. Participants

Table 2 presents the baseline characteristics of the general adult population and individuals diagnosed with AD across six data sources from Croatia, Denmark, Germany, the Netherlands, and the UK. The size of the general adult population ranged from 2,109,064 individuals in IPCI to 29,984,487 in IQVIA DA Germany. Median ages in the general adult population ranged between 40 and 51 years and mean ages between 42 and 50 years. The majority of adults in the general adult population fell within the age category of 18–55 years, accounting for 58.8–78.0% of individuals across data sources. Older age groups represented a small minority: for example, only 1.0–2.7% of the general population were aged 86 years and over. The sex distribution was balanced in the general population with female proportions ranging from approximately 50.3–55.9% across data sources (**Table 2**).

Across the six data sources, after applying the exclusion criteria, the number of individuals diagnosed with AD ranged widely, from 8,908 in IPCI to 110,737 in DK-DHR, with InGef RDB (84,822), CPRD GOLD (66,301), IQVIA DA Germany (52,443), and NAJS (27,795) falling in between (**ANNEX V and Table 2**). Despite differences in population size, the age distribution was consistent across cohorts: the median age at index date fell between 79 and 82 years and the mean age between 78 and 81 years. Most AD cases were in the age group of 76–85 years, representing roughly half of all individuals with AD in each data source. The proportion of individuals aged 86 years and over ranged from 18.9% (NAJS) to 30.7% (CPRD GOLD). A very small proportion of AD cases were younger than 55 years: consistently below 3.0% in all data sources. Sex distributions were also similar across datasets, with females comprising the majority of individuals with AD (approximately 56.2–66.4%) (**Table 2**).

Table 2. Distribution of baseline characteristics among participants number and %, median and IQR, mean and SD) per cohort of interest, by data source.

	Characteristics	NAJS	DK-DHR	InGef RDB	IQVIA DA Germany	IPCI	CPRD GOLD
General adult population	Overall, N	4,143,246	6,664,515	8,836,627	29,984,487	2,109,064	7,617,230
	Median age (IQR)	46 (29–62)	34 (20–53)	41 (27–54)	51 (34–65)	43 (28–59)	42 (28–59)
	Mean age (SD)	46.33 (20.04)	38.55 (19.42)	42.14 (17.96)	50.27 (19.16)	44.56 (19.17)	44.79 (19.61)
	Age groups, in years N (%)						
	Unknown	204 (0.00%)	806 (0.01%)	-	-	-	-
	18–55	2,690,357 (64.93%)	5,198,979 (78.01%)	6,817,836 (77.15%)	17,640,760 (58.83%)	1,475,456 (69.96%)	5,366,590 (70.45%)
	56–65	633,550 (15.29%)	696,951 (10.46%)	995,993 (11.27%)	5,091,097 (16.98%)	276,143 (13.09%)	907,489 (11.91%)
	66–75	441,451 (10.65%)	419,505 (6.29%)	575,452 (6.51%)	3,784,954 (12.62%)	209,813 (9.95%)	708,488 (9.30%)
	76–85	304,925 (7.36%)	265,066 (3.98%)	356,125 (4.03%)	2,803,846 (9.35%)	112,034 (5.31%)	431,423 (5.66%)
	≥86	72,759 (1.76%)	83,208 (1.25%)	91,221 (1.03%)	663,830 (2.21%)	35,618 (1.69%)	203,240 (2.67%)
	Sex, N (%)						
	Female	2,124,165 (51.27%)	3,349,345 (50.26%)	4,365,865 (49.41%)	16,770,623 (55.93%)	1,090,152 (51.69%)	3,914,373 (51.39%)
	Male	2,019,081 (48.73%)	3,315,170 (49.74%)	4,470,762 (50.59%)	13,213,864 (44.07%)	1,018,912 (48.31%)	3,702,857 (48.61%)
Individuals with Alzheimer’s disease	Overall, N	27,795	110,737	84,822	52,443	8,908	66,301
	Median age (IQR) at index date	79 (73–84)	81 (75–85)	82 (76–86)	80 (75–84)	81 (76–86)	82 (77–87)
	Mean age (SD) at index date	77.99 (9.30)	79.74 (8.08)	80.65 (8.38)	78.55 (9.15)	80.22 (8.13)	81.11 (7.95)
	Age groups, in years N (%)						
	Unknown	<5	-	<5	-	-	-
	18–55	632 (2.27%)	1,291 (1.17%)	951 (1.12%)	1,315 (2.51%)	74 (0.83%)	465 (0.70%)

	Characteristics	NAJS	DK-DHR	InGef RDB	IQVIA DA Germany	IPCI	CPRD GOLD
	56–65	1,624 (5.84%)	4,203 (3.80%)	3,491 (4.12%)	2,838 (5.41%)	343 (3.85%)	2,181 (3.29%)
	66–75	6,850 (24.64%)	23,105 (20.86%)	14,132 (16.66%)	10,565 (20.15%)	1,801 (20.22%)	11,369 (17.15%)
	76–85	13,430 (48.32%)	55,478 (50.10%)	42,120 (49.66%)	27,335 (52.12%)	4,281 (48.06%)	31,945 (48.18%)
	≥86	5,255 (18.91%)	26,660 (24.07%)	24,127 (28.44%)	10,390 (19.81%)	2,409 (27.04%)	20,341 (30.68%)
	Sex, N (%)						
	Female	18,443 (66.35%)	67,358 (60.83%)	47,633 (56.16%)	30,935 (58.99%)	5,627 (63.17%)	43,117 (65.03%)
	Male	9,352 (33.65%)	43,380 (39.17%)	37,189 (43.84%)	21,472 (40.94%)	3,281 (36.83%)	23,184 (34.97%)
	Unknown*	-	-	-	36 (0.07%)	-	-

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

IQR=Interquartile Range, SD=Standard Deviation

*Unknown sex is only reported for the individuals with AD and not for the general population. The general population is created by *IncidencePrevalence* and is the denominator population for the incidence/prevalence estimates. Sex is required by *IncidencePrevalence,* meaning individuals with unknown sex are excluded from the denominator population (general population).

9.2. Descriptive data

9.3. Main results

9.3.1. Objective 1 (Incidence and prevalence of AD)

Denominator population for the incidence of AD

For incidence calculations, the total number of individuals that were included after applying the exclusion criterion was 4,143,242 (34,545,813 PY) for NAJS, 6,664,515 (92,763,517.52 PY) for DK-DHR, 8,836,626 (61,762,676 PY) for InGef RDB, 29,978,698 (78,683,903 PY) for IQVIA DA Germany, 2,108,255 (10,838,707 PY) for IPCI, and 7,604,470 (35,240,950 PY) for CPRD GOLD (**Table 3**).

Overall incidence of AD

Across the six data sources, the overall incidence (95% CI) per 100,000 PY for both sexes and all age groups, over the total study period (2014–2024), showed considerable variation, ranging from 53.85 in IQVIA DA Germany to 137.33 in InGef RDB. For each specific data source, the incidence was 80.45 (79.50–81.40) for NAJS, 119.38 (118.67–120.08) for DK-DHR, 137.33 (136.41–138.26) for InGef RDB, 53.85 (53.34–54.37) for IQVIA DA Germany, 64.95 (63.44–66.49) for IPCI, and 99.70 (98.67–100.75) for CPRD GOLD (**Table 3**).

Table 3. Overall incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, 2014–2024.

Data sources	Number of participants	Follow-up (PYs)	Number of individuals with AD	Incidence Rates/100,000 PYs (95% CIs)
NAJS	4,143,242	34,545,813.32	27,791	80.45 (79.50–81.40)
DK-DHR	6,664,515	92,763,517.52	110,737	119.38 (118.67–120.08)
InGef RDB	8,836,626	61,762,676.33	84,821	137.33 (136.41–138.26)
IQVIA DA Germany	29,978,698	78,683,903.01	42,374	53.85 (53.34–54.37)
IPCI	2,108,255	10,838,707.16	7,040	64.95 (63.44–66.49)
CPRD GOLD	7,604,470	35,240,950.08	35,137	99.70 (98.67–100.75)

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
AD=Alzheimer's disease, PY=Person-years, CI=Confidence Interval

Incidence of AD over time

There was no clear trend (increase or decrease) in the incidence of AD over the years across the data sources. Some data sources (mainly NAJS and DK-DHR) generally showed stable incidence rates, while InGef RDB showed a decrease over the years 2016–2020 which stabilised in the years thereafter (2020–2023), and IQVIA DA Germany showed stable incidence estimates from 2014–2021 and an increase thereafter (**Figure 6 and ANNEX VI Table S1–S6**). The incidence (95% CI) per 100,000 PY ranged from 89.38 (86.23–92.62) in 2018 to 81.65 (78.71–84.67) in 2024 for NAJS, for DK-DHR from 121.66 (118.43–124.95) in 2014 to 144.77 (141.37–148.23) in 2024, for InGef RDB from 198.41 (194.88–201.98) in 2016 to 143.06 (140.14–146.04) in 2023, for IQVIA DA Germany from 41.51 (39.90–43.16) in 2014 to 98.75 (96.09–101.46) in 2024, for IPCI from 65.97 (60.25–72.09) in 2014 to 45.68 (41.71–49.92) in 2024, and for CPRD GOLD from 89.35 (86.69–92.06) in 2014 to 120.16 (115.70–124.74) in 2024 (**Figure 6 and ANNEX VI Table S1–S6**).

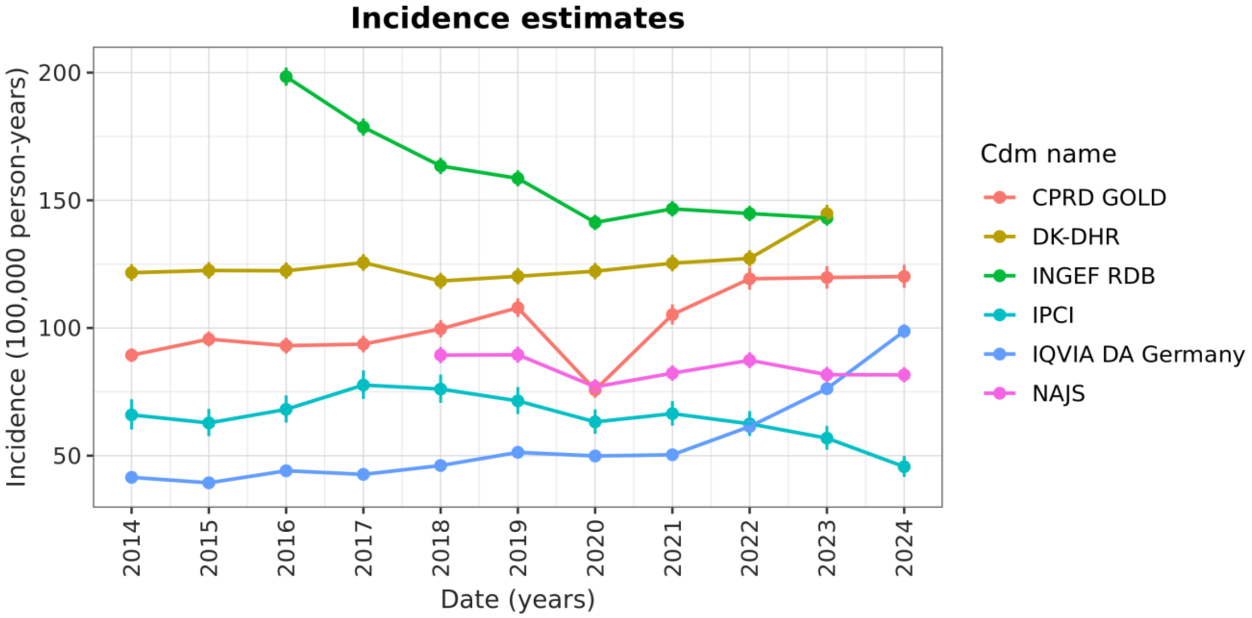

Figure 6. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall incidence of AD by sex

The overall incidence of AD over the study period (2014–2024) was higher among females than males across all data sources, with an incidence (95% CI) per 100,000 PY observed of 100.99 (99.54–102.46) for females and 57.42 (56.26–58.59) for males in NAJS, 143.08 (142.00–144.16) for females and 94.95 (94.06–95.85) for males in DK-DHR, 153.08 (151.71–154.46) for females and 121.35 (120.12–122.59) for males in InGef RDB, 55.33 (54.65–56.02) for females and 51.84 (51.06–52.62) for males in IQVIA DA Germany, 78.51 (76.20–80.88) for females and 50.61 (48.71–52.57) for males in IPCI, and 124.95 (123.32–126.60) for females and 73.68 (72.41–74.97) for males in CPRD GOLD (**ANNEX VI Table S1–S6**).

Incidence of AD over time by sex

The distribution of incidence over time shown in **Figure 7** was similar between females and males and similar to the overall incidence trend shown in **Figure 6**. A higher incidence among females than males was also seen among most age groups (**ANNEX VI Table S1–S6**), with some exceptions, mainly in the younger age groups.

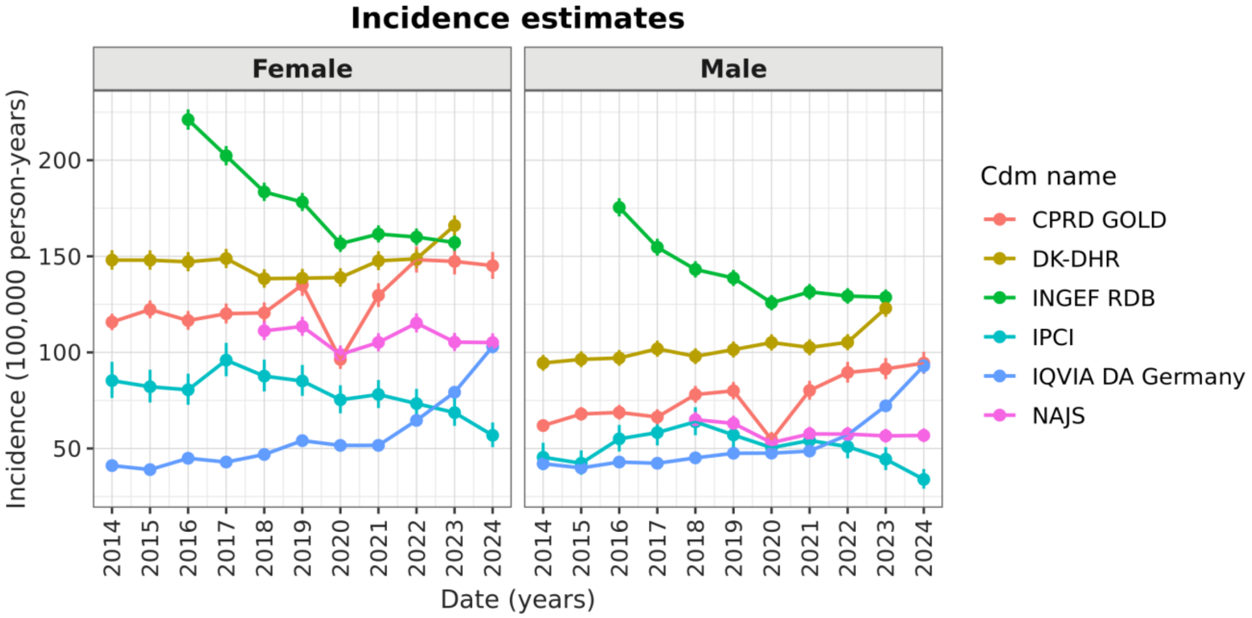

Figure 7. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, stratified by sex, 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall Incidence of AD by age

The overall incidence (95% CI) over the study period (2014–2024) showed a much higher incidence for late-onset AD (≥66 years) compared to early-onset AD (18–65 years) across all data sources. Early-onset incidence per 100,000 PY ranged from 3.71 in IPCI to 8.72 in InGef RDB, while late-onset incidence ranged from 161.14 in IQVIA DA Germany to 742.40 in InGef RDB. By data source, incidence for early-onset versus late-onset AD was 8.65 (8.30–9.02) versus 301.48 (297.80–305.20) in NAJS, 7.50 (7.31–7.71) versus 538.13 (534.88–541.39) in DK-DHR, 8.72 (8.47–8.98) versus 742.40 (737.28–747.55) for InGef RDB, 5.52 (5.32–5.72) versus 161.14 (159.55–162.74) in IQVIA DA Germany, 3.71 (3.31–4.15) versus 273.96 (267.45–280.59) in IPCI, and 4.48 (4.24–4.74) versus 445.73 (441.00–450.50) in CPRD GOLD (**ANNEX VI Table S1–S6**).

Incidence of AD over time by age

The distribution of incidence over time shown in **Figure 8** for late-onset AD was similar to the overall incidence trend shown in **Figure 6**. The incidence (95% CI) per 100,000 PY of early-onset AD remained consistently low over the years across all data sources, below 12.24 (11.30–13.24). In contract, late-onset AD showed a much higher incidence 100,000 PY, with values ranging from 347.07 (334.22–360.28) in 2018 to 298.65 (287.54–310.08) in 2024 for NAJS, from 539.30 (524.62–554.28) in 2014 to 587.24 (573.19–601.56) in 2023 for DK-DHR, from 1,146.32 (1,125.42–1,167.50) in 2016 to 718.21 (703.03–733.63) in 2023 for InGef RDB, from 131.11 (125.77–136.62) in 2014 to 263.72 (256.37–271.22) in 2024 for IQVIA DA Germany, from 298.16 (271.56–326.66) in 2014 to 182.58 (166.42–199.90) in 2024 for IPCI, and from 410.96 (398.51–423.70) in 2014 to 507.64 (488.55–527.28) in 2024 for CPRD GOLD (**Figure 8**).

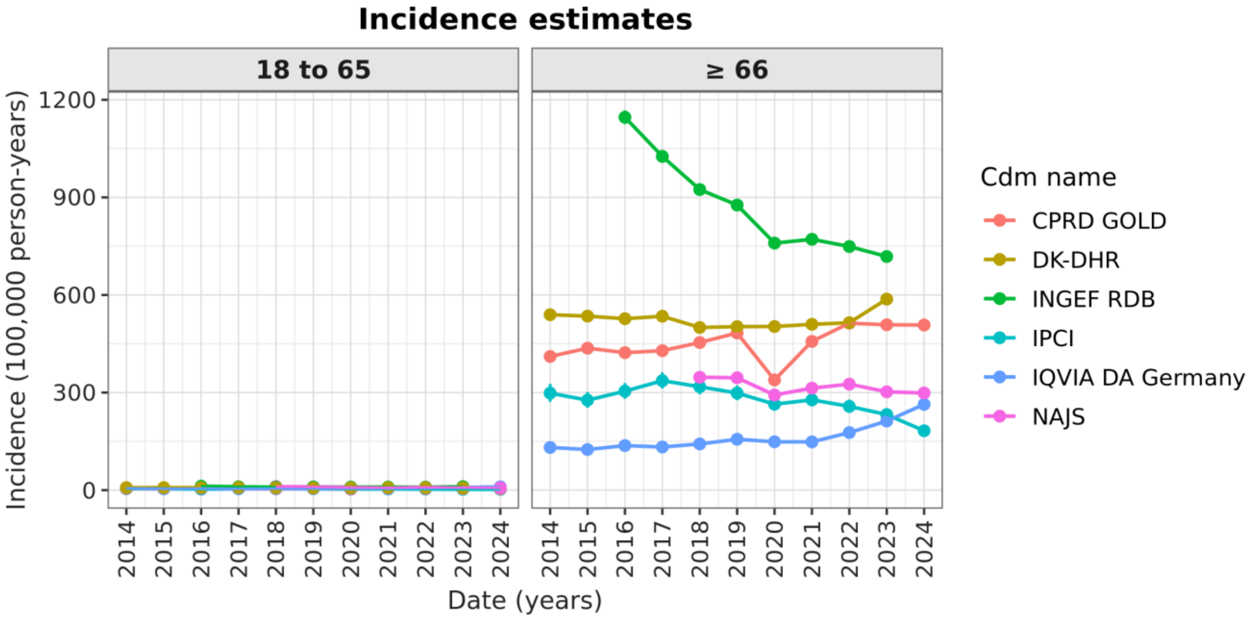

Figure 8. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, stratified by early-onset AD (18–65 years) and late-onset AD (≥66 years), 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall incidence of AD by age groups

When stratifying the results into multiple age categories, the overall incidence over the study period (2014–2024) showed an increase with increasing age across all data sources. A significantly higher incidence was observed among the two oldest age groups (76–85 and 86 years and over), compared to the younger age groups (18–55, 56–65, 66–75 years), among all data sources (**ANNEX VI Table S1–S6**).

Incidence of AD over time by age groups

In NAJS, the incidence (95% CI) per 100,000 PY ranged from 4.62 (3.73–5.66) in 2018 to 2.23 (1.63–2.98) in 2024 among 18–55 year olds, from 32.32 (27.90–37.25) in 2018 to 23.77 (20.00–28.05) in 2024 among 56–65 year olds, from 160.18 (148.36–172.70) in 2018 to 139.51 (129.57–150.01) in 2024 among 66–75 year olds, from 529.90 (503.58–557.24) in 2018 to 506.70 (480.55–533.90) in 2024 among 76–85 year olds, and from 671.02 (616.39–729.18) in 2018 to 528.66 (486.79–573.18) in 2024 among individuals aged 86 years and over (**Figure 9**).

In DK-DHR, the incidence (95% CI) per 100,000 PY ranged from 2.47 (1.92–3.13) in 2014 to 1.75 (1.30–2.31) in 2023 among 18–55 year olds, from 28.86 (24.98–33.17) in 2014 to 27.23 (23.66–31.18) in 2023 among 56–65 year olds, from 181.08 (170.37–192.29) in 2014 to 239.20 (227.09–251.78) in 2023 among 66–75 year olds, from 947.16 (910.80–984.60) in 2014 to 894.24 (865.61–923.57) in 2023 among 76–85 year olds, and from 1,654.39 (1,571.58–1,740.43) in 2014 to 1,408.45 (1,337.46–1,482.23) in 2023 among individuals aged 86 years and over (**Figure 10**).

In InGef RDB, the incidence (95% CI) per 100,000 PY ranged from 3.63 (3.08–4.26) in 2016 to 2.77 (2.27–3.34) in 2023 among 18–55 year olds, from 51.88 (47.30–56.78) in 2016 to 35.61 (32.37–39.09) in 2023 among 56–65 year olds, from 385.21 (368.84–402.12) in 2016 to 228.85 (217.75–240.36) in 2023 among 66–75 year olds, from 1,627.39 (1,586.31–1,669.27) in 2016 to 1,147.02 (1,112.81–1,182.02) in 2023 among 76–85 year olds, and from 3,680.42 (3,558.13–3,805.84) in 2016 to 2,259.64 (2,174.67–2,347.08) in 2023 among individuals aged 86 years and over (**Figure 11**).

In IQVIA DA Germany, the incidence (95% CI) per 100,000 PY ranged from 2.93 (2.37–3.59) in 2014 to 3.23 (2.54–4.05) in 2024 among 18–55 year olds, from 13.92 (11.83–16.28) in 2014 to 24.36 (21.56–27.42) in 2024 among 56–65 year olds, from 62.10 (57.32–67.17) in 2014 to 97.12 (90.89–103.68) in 2024 among 66–75 year olds, from 202.38 (191.61–213.61) in 2014 to 371.32 (356.88–386.20) in 2024 among 76–85 year olds, and from 450.06 (402.31–501.92) in 2014 to 590.20 (560.60–620.96) in 2024 among individuals aged 86 years and over (**Figure 12**).

In IPCI, the incidence (95% CI) per 100,000 PY ranged from 1.30 (0.52–2.68) in 2015 to 0.96 (0.35–2.08) in 2024 among 18–55 year olds, from 19.24 (12.20–28.87) in 2014 to 7.14 (3.80–12.21) in 2024 among 56–65 year olds, from 106.90 (86.79–130.28) in 2014 to 56.32 (44.66–70.09) in 2024 among 66–75 year olds, from 508.03 (446.40–575.79) in 2014 to 245.35 (214.14–279.84) in 2024 among 76–85 year olds, and from 801.78 (663.66–960.18) in 2014 to 698.03 (595.88–812.66) in 2024 among individuals aged 86 years and over (**Figure 13**).

In CPRD GOLD, the incidence (95% CI) per 100,000 PY ranged from 1.25 (0.89–1.71) in 2014 to 1.09 (0.61–1.79) in 2024 among 18–55 year olds, from 19.29 (16.15–22.86) in 2014 to 16.60 (12.75–21.24) in 2024 among 56–65 year olds, from 125.02 (115.90–134.66) in 2014 to 145.56 (131.87–160.28) in 2024 among 66–75 year olds, from 625.62 (598.61–653.54) in 2014 to 723.34 (684.63–763.66) in 2024 among 76–85 year olds, and from 1,136.53 (1,077.97–1,197.44) in 2014 to 1,571.26 (1,472.71–1,674.66) in 2024 among individuals aged 86 years and over (**Figure 14**).

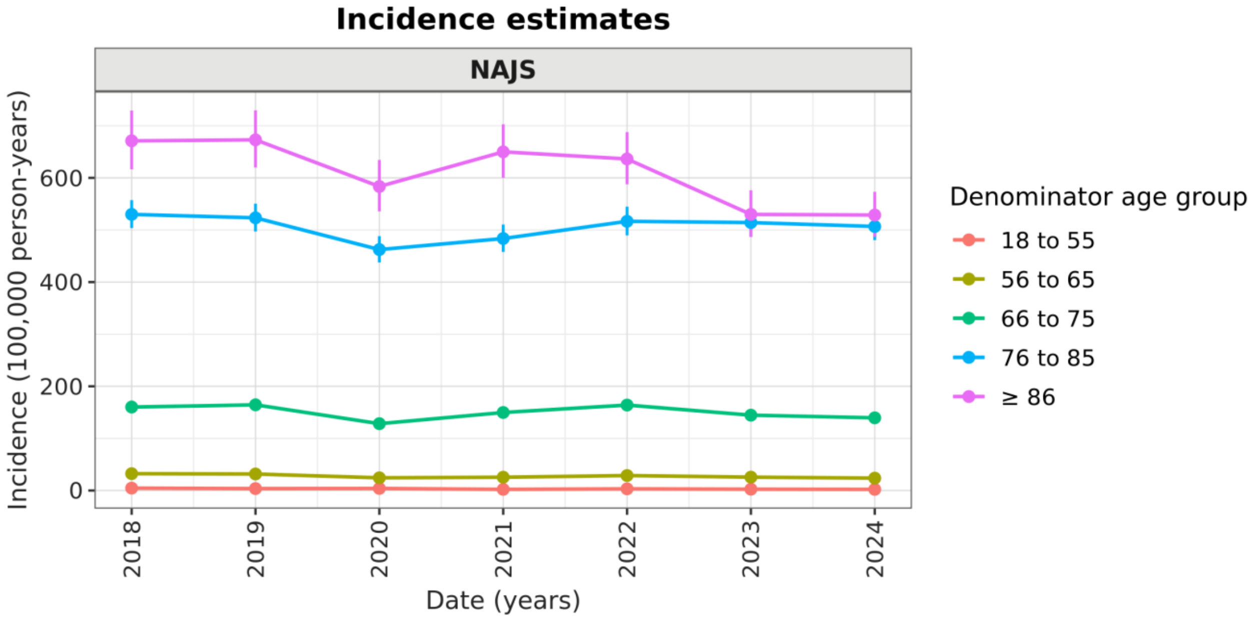

Figure 9. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in NAJS, stratified by age categories, 2018–2024.

NAJS=National Public Health Information System

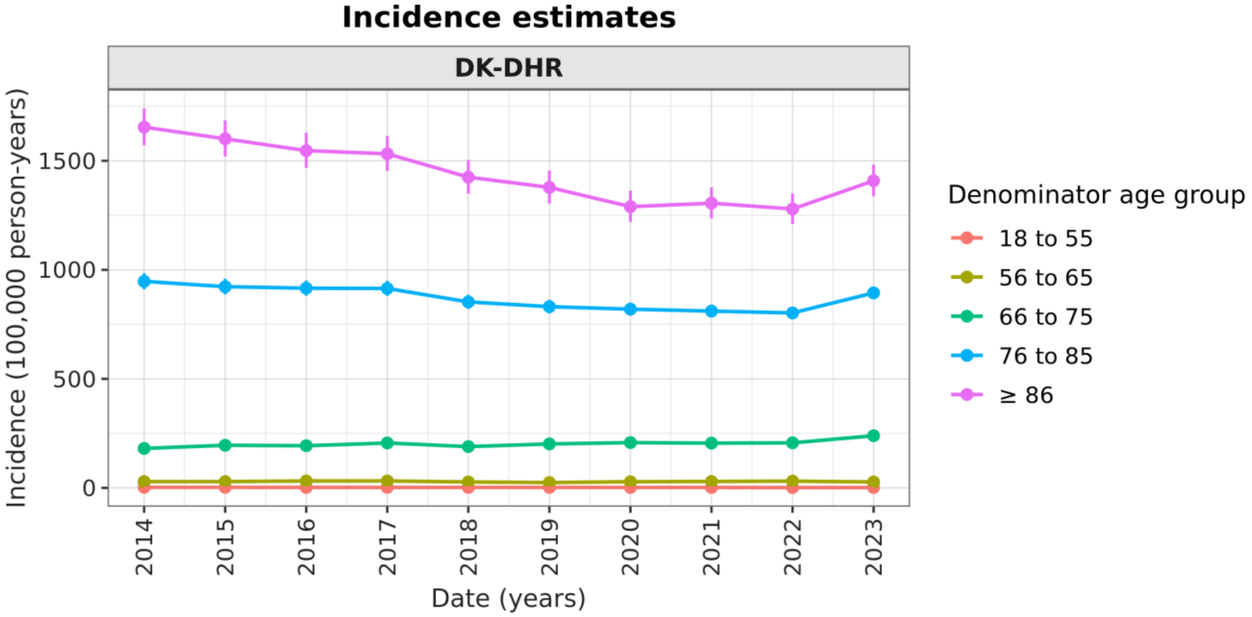

Figure 10. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in DK-DHR, stratified by age categories, 2014–2023.

DK-DHR=Danish Data Health Registries

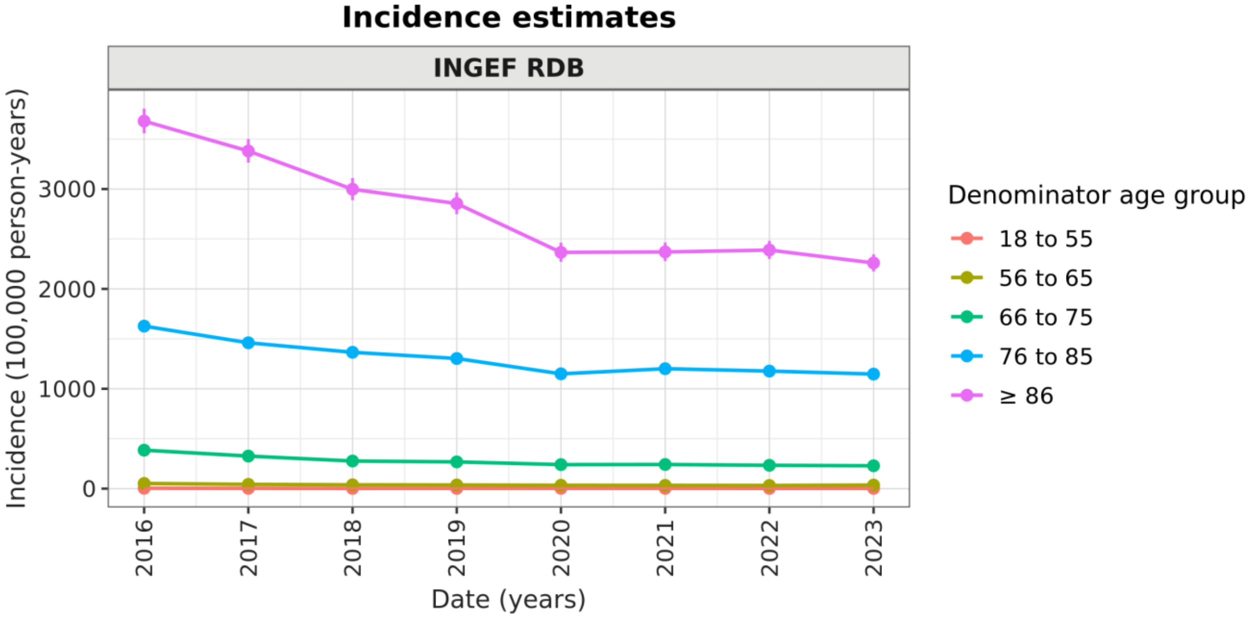

Figure 11. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in InGef RDB, stratified by age categories, 2016–2023.

InGef RDB=InGef Research Database

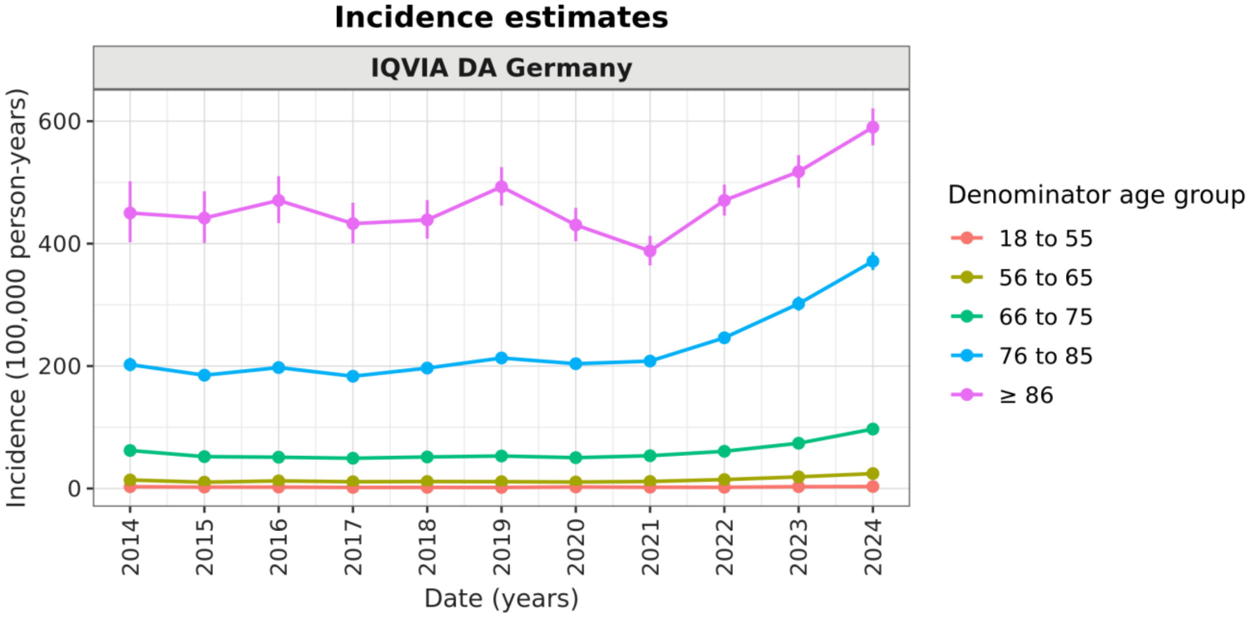

Figure 12. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in IQVIA DA Germany, stratified by age categories, 2014–2024.

IQVIA DA Germany=IQVIA Disease Analyzer Germany

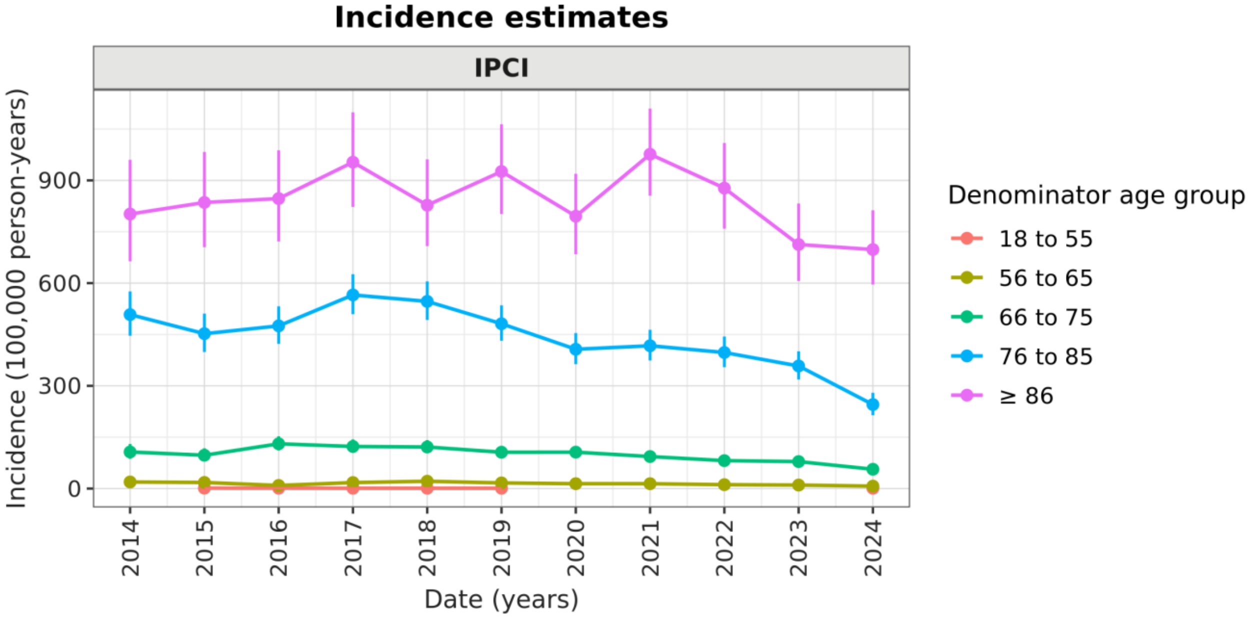

Figure 13. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in IPCI, stratified by age categories, 2014–2024.

IPCI=Integrated Primary Care Information

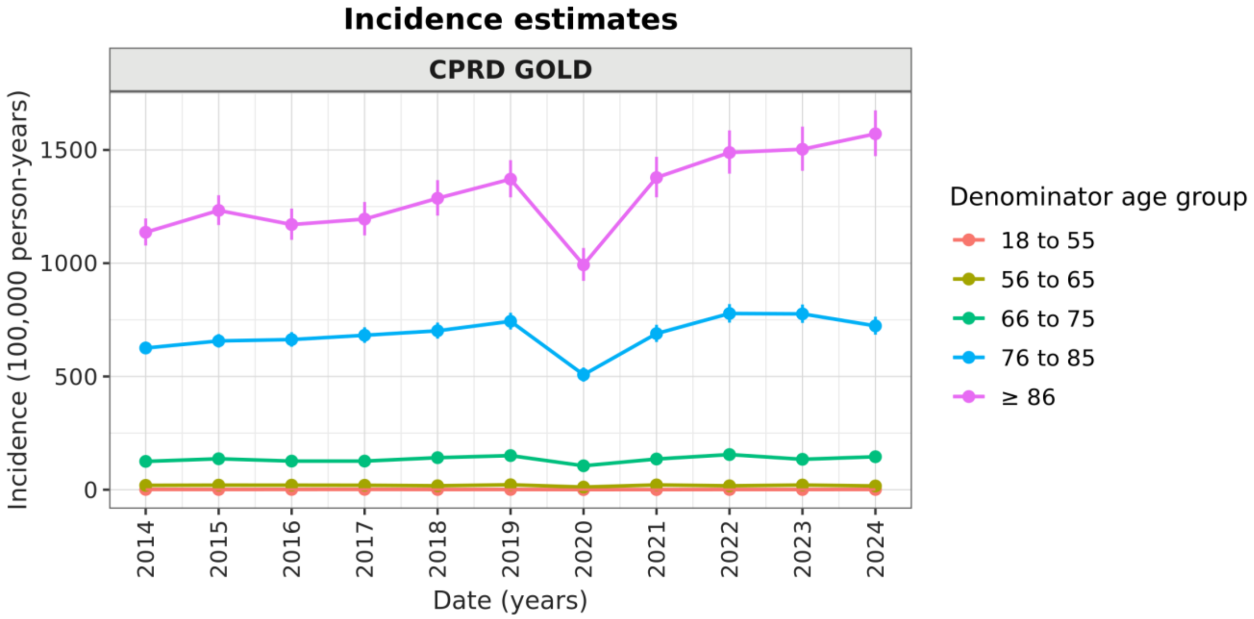

Figure 14. Yearly incidence rates per 100,000 PYs (95% CIs) of AD in CPRD GOLD, stratified by age categories, 2014–2024.

CPRD=Clinical Practice Research Datalink

Denominator population for the prevalence of AD

For prevalence calculations, the total number of individuals included was 4,143,246 for NAJS, 6,664,515 for DK-DHR, 8,836,627 for InGef RDB, 29,984,487 for IQVIA DA Germany, 2,109,064 for IPCI, and 7,617,230 for CPRD GOLD (**Table 4**).

Overall prevalence of AD

The overall prevalence (95% CI) for both sexes and all age groups was 0.67% (0.66–0.68%) for NAJS, 1.66% (1.65–1.67%) for DK-DHR, 0.96% (0.95–0.97%) for InGef RDB, 0.16% (0.16–0.16%) for IQVIA DA Germany, 0.37% (0.36–0.38%) for IPCI, and 0.63% (0.62–0.63%) for CPRD GOLD (**Table 4**).

Table 4. Overall prevalence (95% CI) of AD in the general adult population in NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, 2014–2024.

Data source	Number of participants	Number of individuals with AD	Prevalence (95% CI)
NAJS	4,143,246	27,795	0.67 (0.66–0.68)
DK-DHR	6,664,515	110,737	1.66 (1.65–1.67)
InGef RDB	8,836,627	84,822	0.96 (0.95–0.97)
IQVIA DA Germany	29,984,487	48,160	0.16 (0.16–0.16)
IPCI	2,109,064	7,849	0.37 (0.36–0.38)
CPRD GOLD	7,617,230	47,897	0.63 (0.62–0.63)

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
AD=Alzheimer's disease, CI=Confidence Interval

Prevalence of AD over time

The prevalence of AD showed a higher increase over the years for NAJS and InGef RDB and was more stable for the remaining data sources (DK-DHR, IQVIA DA Germany, IPCI, and CPRD GOLD) (**Figure 15**). By data source, the prevalence (95% CI) ranged from 0.24% (0.23–0.24%) in 2018 to 0.37% (0.37–0.38%) in 2024 for NAJS, from 0.62% (0.61–0.62%) in 2014 to 0.69 (0.69–0.70%) in 2023 for DK-DHR, from 0.19% (0.19–0.19%) in 2016 to 0.66% (0.65–0.66%) in 2023 for InGef RDB, from 0.11% (0.11–0.11%) in 2014 to 0.23% (0.22–0.23%) in 2024 for IQVIA DA Germany, from 0.15% (0.14–0.16%) in 2014 to 0.20% (0.19–0.20%) in 2024 for IPCI, and from 0.32% (0.31–0.32%) in 2014 to 0.46% (0.45–0.47%) in 2024 for CPRD GOLD. (**Figure 15 and ANNEX VII Table S1–S6**).

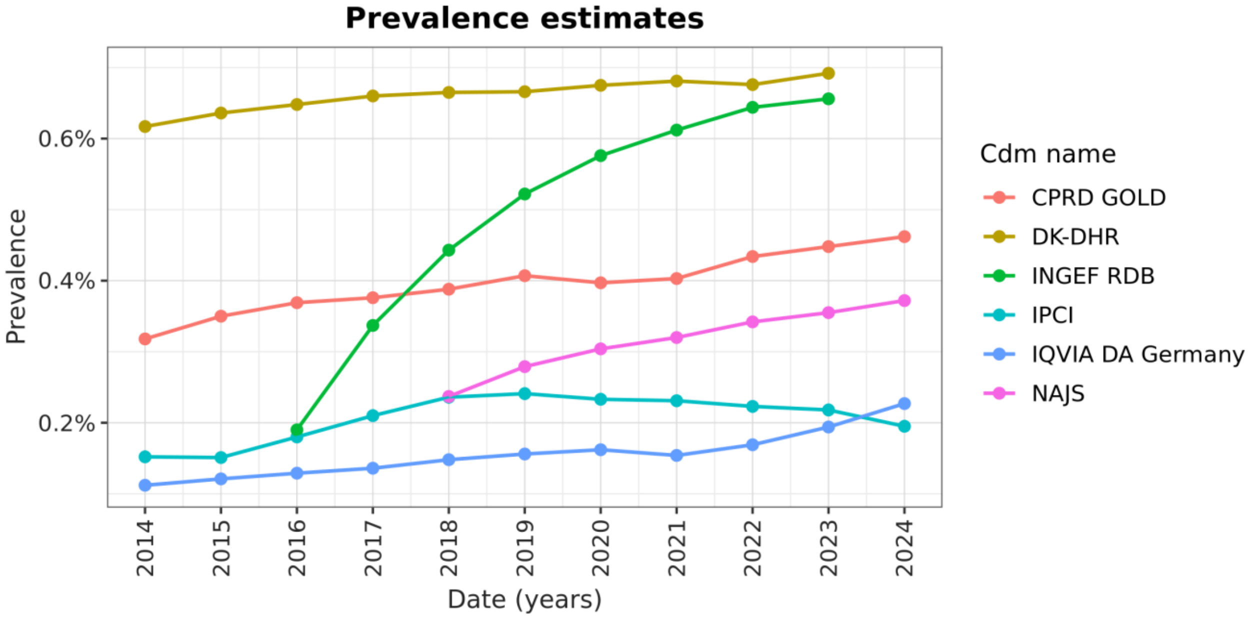

Figure 15. Overall prevalence (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall prevalence of AD by sex

The overall prevalence (95% CI) of AD over the study period (2014–2024) was higher among females than males among all data sources, with a prevalence of 0.87% (0.86–0.88%) for females and 0.46% (0.45–0.47%) for males for NAJS, 2.01% (2.00–2.03%) for females and 1.31% (1.30–1.32%) for males in DK-DHR, 1.09% (1.08–1.10%) for females and 0.83% (0.82–0.84%) for males in InGef RDB, 0.17% (0.17–0.17%) for females and 0.15% (0.15–0.15%) for males in IQVIA DA Germany, 0.45% (0.44–0.46%) for females and 0.29% (0.28–0.30%) for males in IPCI, and 0.79% (0.78–0.80%) for females and 0.46% (0.45–0.47%) for males in CPRD GOLD (**ANNEX VII Table S1–S6**).

Prevalence of AD over time by sex

The distribution of prevalence over time shown in **Figure 16** was similar between females and males and similar to the overall prevalence trend shown in **Figure 15**. The prevalence (95% CI) of AD was higher among females than males among all data sources. For NAJS, the prevalence ranged from 0.30% (0.29–0.31%) in 2018 to 0.50% (0.49–0.51%) in 2024 for females and from 0.17% (0.16–0.17%) in 2018 to 0.24% (0.24–0.25%) in 2024 for males. For DK-DHR, the prevalence ranged from 0.78% (0.77–0.79%) in 2014 to 0.83% (0.82–0.84%) in 2023 for females and from 0.45% (0.44–0.46%) in 2014 to 0.55% (0.54–0.56%) in

2023 for males. For InGef RDB, the prevalence ranged from 0.21% (0.21–0.22%) in 2016 to 0.75% (0.74–0.76%) in 2023 for females and from 0.17% (0.16–0.17%) in 2016 to 0.56% (0.55–0.57%) in 2023 for males. For IQVIA DA Germany, the prevalence ranged from 0.11% (0.11–0.12%) in 2014 to 0.24% (0.24–0.25%) in 2024 for females and from 0.11% (0.11–0.12%) in 2014 to 0.21% (0.21–0.22%) in 2024 for males. For IPCI, the prevalence ranged from 0.20% (0.18–0.21%) in 2014 to 0.23% (0.22–0.24%) in 2024 for females and from 0.10% (0.09–0.11%) in 2014 to 0.16% (0.15–0.17%) in 2024 for males. For CPRD GOLD, the prevalence ranged from 0.41% (0.40–0.42%) in 2014 to 0.57% (0.56–0.58%) in 2024 for females and from 0.22% (0.21–0.23%) in 2014 to 0.35% (0.34–0.36%) in 2024 for males (**Figure 15**).

A higher prevalence among females than males seen among most age groups, with some exceptions, mainly in the younger age groups (**ANNEX VII Table S1–S6**).

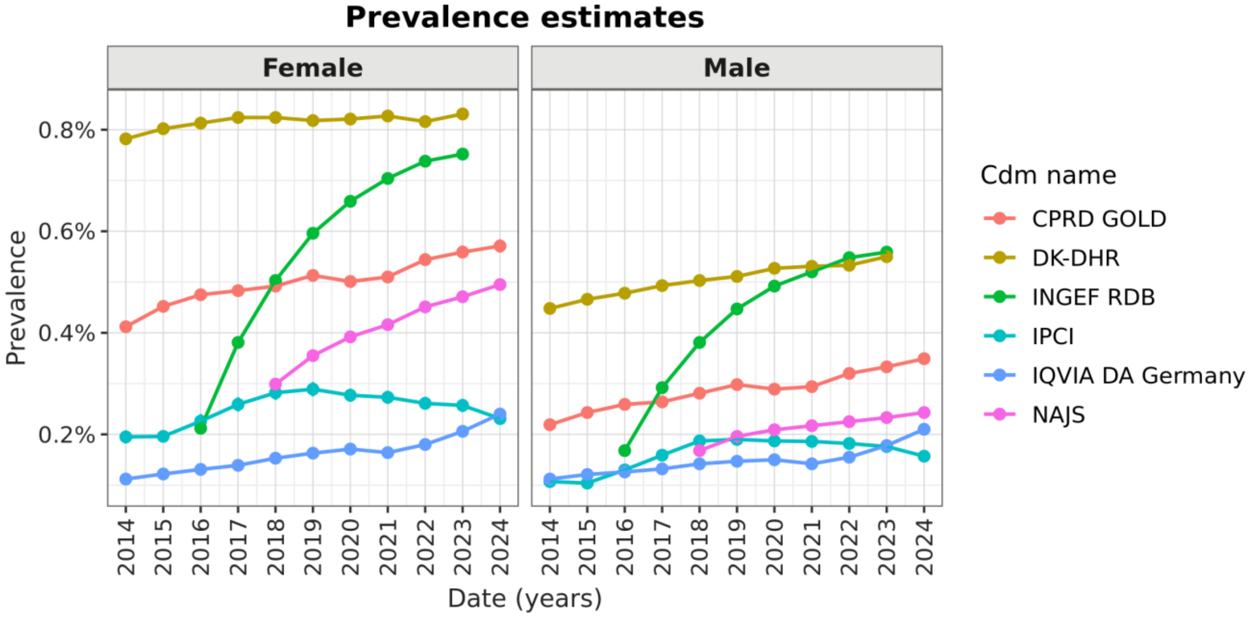

Figure 16. Prevalence (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, stratified by sex, 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall prevalence of AD by age

The overall prevalence (95% CI) of AD over the study period (2014–2024) was higher for late-onset AD (≥66 years) compared to early-onset AD (18–65 years) across all data sources. Early-onset prevalence ranged from 0.02% in IQVIA DA Germany and IPCI to 0.09% in DK-DHR. By data source, prevalence for early-onset versus late-onset was 0.07% (0.06–0.07%) versus 1.90% (1.87–1.92%) for NAJS, 0.09% (0.09–0.10%) versus 5.09% (5.06–5.12%) for DK-DHR, 0.06% (0.06–0.06%) versus 4.54% (4.52–4.58%) for InGef RDB, 0.02% (0.01–0.02%) versus 0.53% (0.53–0.54%) for IQVIA DA Germany, 0.02% (0.02–0.02%) versus 1.47% (1.44–1.51%) in IPCI, and 0.03% (0.03–0.03%) versus 2.65% (2.62–2.67%) in CPRD GOLD (**ANNEX VII Table S1–S6**).

Prevalence of AD over time by age

Over the years, the prevalence of early-onset AD (18–65 years) was below 0.05% among all data sources. For late-onset AD, the prevalence showed a similar trend as the overall prevalence estimates in **Figure 15**. By data source, the prevalence (95% CI) ranged from 0.88% (0.86–0.90%) in 2018 to 1.28% (1.26–1.31%) in 2024 in NAJS, from 2.62% (2.59–2.65%) in 2014 to 2.69% (2.66–2.72%) in 2023 in DK-DHR, from 1.09%

(1.07–1.11%) in 2016 to 3.21% (3.18–3.24%) in 2023 in InGef RDB, from 0.36% (0.36–0.37%) in 2014 to 0.65% (0.64–0.66%) in 2024 in IQVIA DA Germany, from 0.66% (0.63–0.70%) in 2014 to 0.77% (0.74–0.80%) in 2024 in IPCI, and ranged from 1.47% (1.45–1.49%) in 2014 to 1.96% (1.92–1.99%) in 2024 in CPRD GOLD (**Figure 17**).

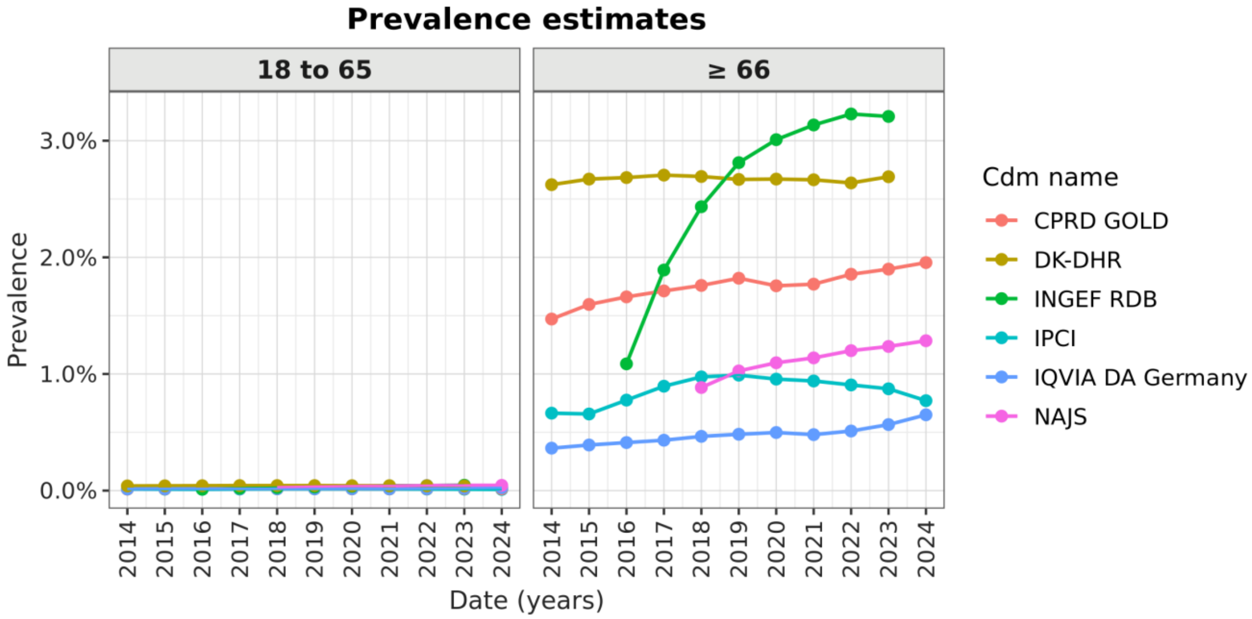

Figure 17. Prevalence (95% CIs) of AD in data sources NAJS, DK-DHR, InGef RDB, IQVIA DA Germany, IPCI, CPRD GOLD, stratified by early-onset AD (18–65 years) and late-onset AD (≥66 years), 2014–2024.

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink

Overall prevalence of AD by age groups

When stratifying the results into multiple age categories, the overall prevalence over the whole study period (2014–2024) showed an increase of the prevalence of AD with increasing age across all data sources. A significantly higher prevalence was observed among the two oldest age groups (76–85 and 86 years and over), compared to the younger age groups (18–55, 56–65, 66–75 years), among all data sources (**ANNEX VII Table S1–S6**).

Prevalence of AD over time by age groups

In NAJS, the prevalence (95% CI) ranged from 0.01% (0.01–0.01%) in 2018 to 0.02% (0.02–0.02%) in 2024 among 18–55 year olds, from 0.09% (0.08–0.09%) in 2018 to 0.13% (0.12–0.14%) in 2024 among 56–65 year olds, from 0.43% (0.42–0.45%) in 2018 to 0.60% (0.58–0.62%) in 2024 among 66–75 year olds, from 1.33% (1.29–1.37%) in 2018 to 2.12% (2.07–2.17%) in 2024 among 76–85 year olds, and from 1.61% (1.54–1.69%) in 2018 to 2.44% (2.36–2.52%) in 2024 among individuals aged 86 years and over (**Figure 18**).

In DK-DHR, the prevalence (95% CI) ranged from 0.01% (0.01–0.01%) in 2014 to 0.00% (0.00–0.00%) in 2023 among 18–55 year olds, from 0.10% (0.10–0.20%) in 2014 to 0.15% (0.14–0.16%) in 2023 among 56–65 year olds, from 0.78% (0.75–0.80%) in 2014 to 0.91% (0.89–0.93%) in 2023 among 66–75 year olds, from 4.22% (4.16–4.30%) in 2014 to 3.75% (3.70–3.81%) in 2023 among 76–85 year olds, and from 9.01% (8.93–9.26%) in 2014 to 8.19% (8.05–8.34%) in 2023 among individuals aged 86 years and over(**Figure 19**).

In InGef RDB, the prevalence (95% CI) ranged from 0.00% (0.00–0.00%) in 2016 to 0.01% (0.01–0.01%) in 2023 among 18–55 year olds, from 0.05% (0.04–0.05%) in 2016 to 0.15% (0.15–0.16%) in 2023 among 56–

65 year olds, from 0.35% (0.34–0.37%) in 2016 to 0.91% (0.89–0.93%) in 2023 among 66–75 year olds, from 1.51% (1.47–1.55%) in 2016 to 4.96% (4.89–5.02%) in 2023 among 76–85 year olds, and from 3.32% (3.32–3.34%) in 2016 to 10.15% (10.00–10.30%) in 2023 among individuals aged 86 years and over (**Figure 20**).

In IQVIA DA Germany, the prevalence (95% CI) was 0.00% (0.00–0.00%) among 18–55 year olds over the years 2014–2024, ranged from 0.00% (0.00–0.00%) in 2014 to 0.06% (0.06–0.06%) in 2024 among 56–65 year olds, from 0.15% (0.14–0.15%) in 2014 to 0.21% (0.21–0.22%) in 2024 among 66–75 year olds, from 0.58% (0.56–0.58%) in 2014 to 0.87% (0.85–0.89%) in 2024 among 76–85 year olds, and from 1.29% (1.22–1.36%) in 2014 to 1.64% (1.60–1.68%) in 2024 among individuals aged 86 years and over (**Figure 21**).

In IPCI, the prevalence (95% CI) was 0.00% (0.00–0.00%) among 18–55 year olds over the years 2014–2024, ranged from 0.01% (0.00–0.01%) in 2014 to 0.04% (0.03–0.05%) in 2024 among 56–65 year olds, from 0.25% (0.22–0.28%) in 2014 to 0.26% (0.24–0.29%) in 2024 among 66–75 year olds, from 1.10% (1.02–1.19%) in 2014 to 1.10% (1.04–1.17%) in 2024 among 76–85 year olds, and from 1.79% (1.62-1.99%) in 2014 to 2.37% (2.20–2.55%) in 2024 among individuals aged 86 years and over (**Figure 22**).

In CPRD GOLD, the prevalence (95% CI) was 0.00% (0.00–0.00%) among 18–55 year olds over the years 2014–2024, ranged from 0.06% (0.06–0.07%) in 2014 to 0.08% (0.07–0.08%) in 2024 among 56–65 year olds, from 0.42% (0.40–0.44%) in 2014 to 0.50% (0.47–0.51%) in 2024 among 66–75 year olds, from 2.15% (2.10–2.19%) in 2014 to 2.68% (2.61–2.75%) in 2024 among 76–85 year olds, and from 4.04% (3.94–4.13%) in 2014 to 6.01% (5.84–6.18%) in 2024 among individuals aged 86 years and over (**Figure 23**).

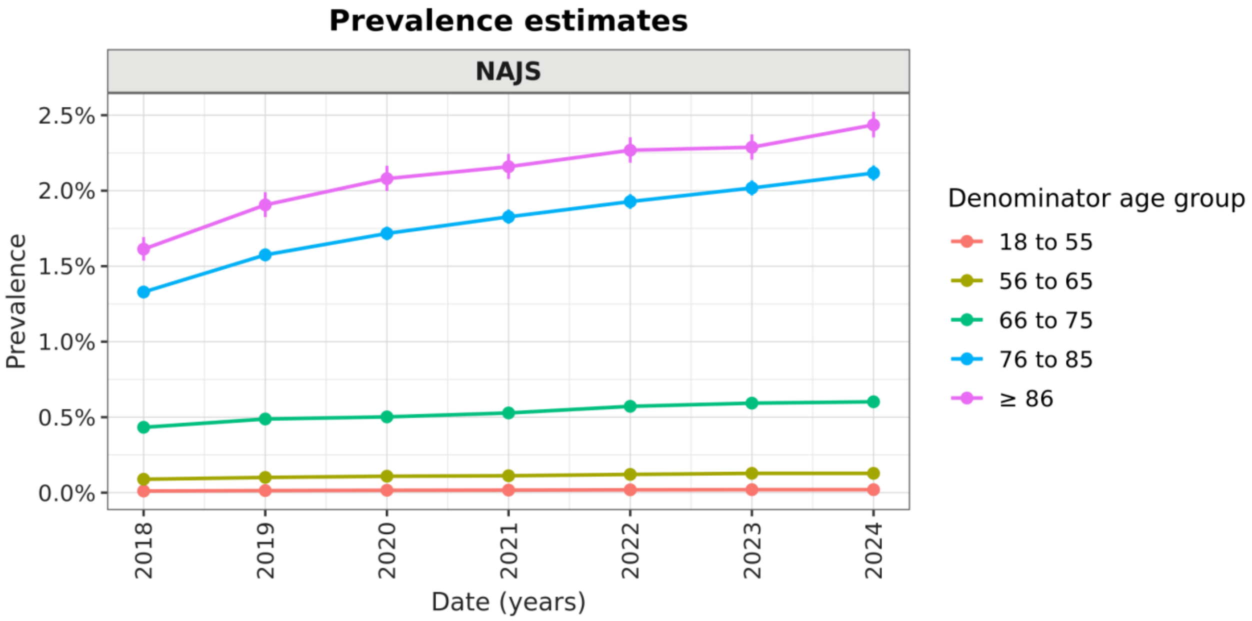

Figure 18. Prevalence (95% CIs) of AD in NAJS, stratified by age categories, 2018–2024.

NAJS=National Public Health Information System

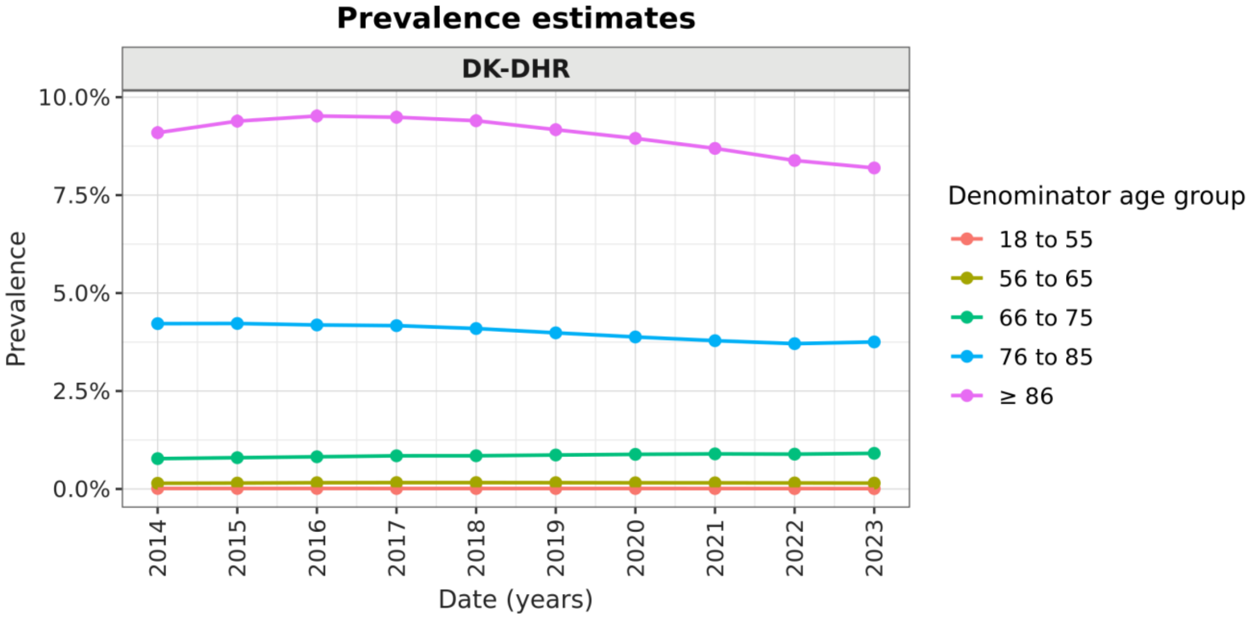

Figure 19. Prevalence (95% CIs) of AD in DK-DHR, stratified by age categories, 2014–2023.

DK-DHR=Danish Data Health Registries

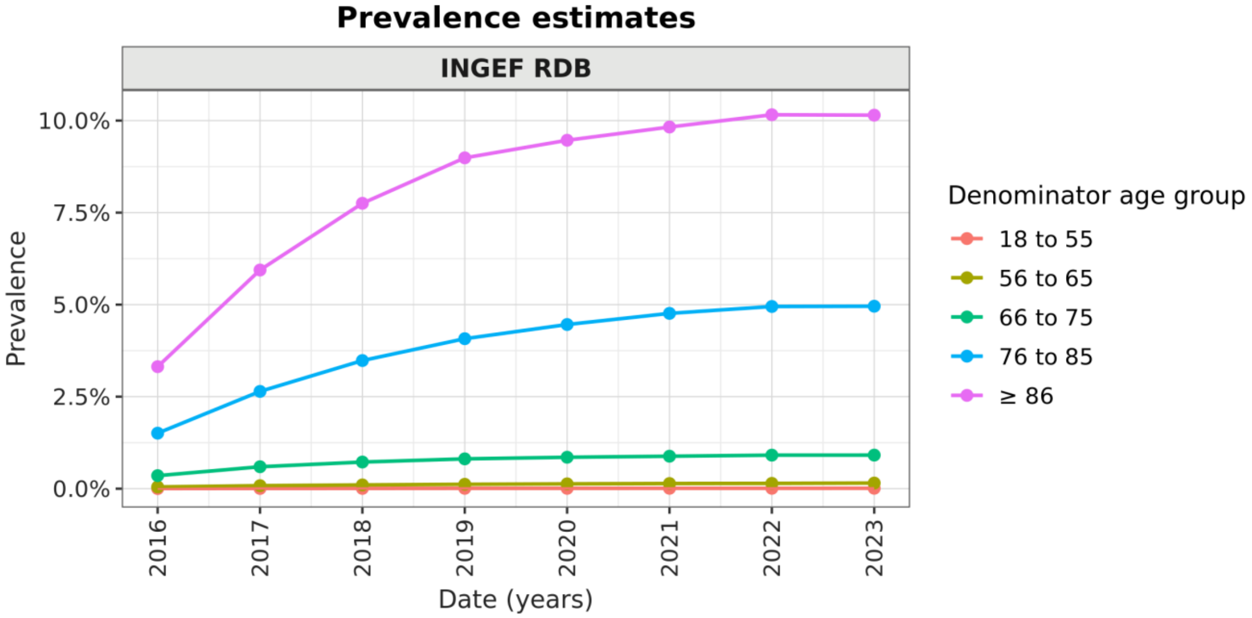

Figure 20. Prevalence (95% CIs) of AD in InGef RDB, stratified by age categories, 2016–2023.

InGef RDB=InGef Research Database

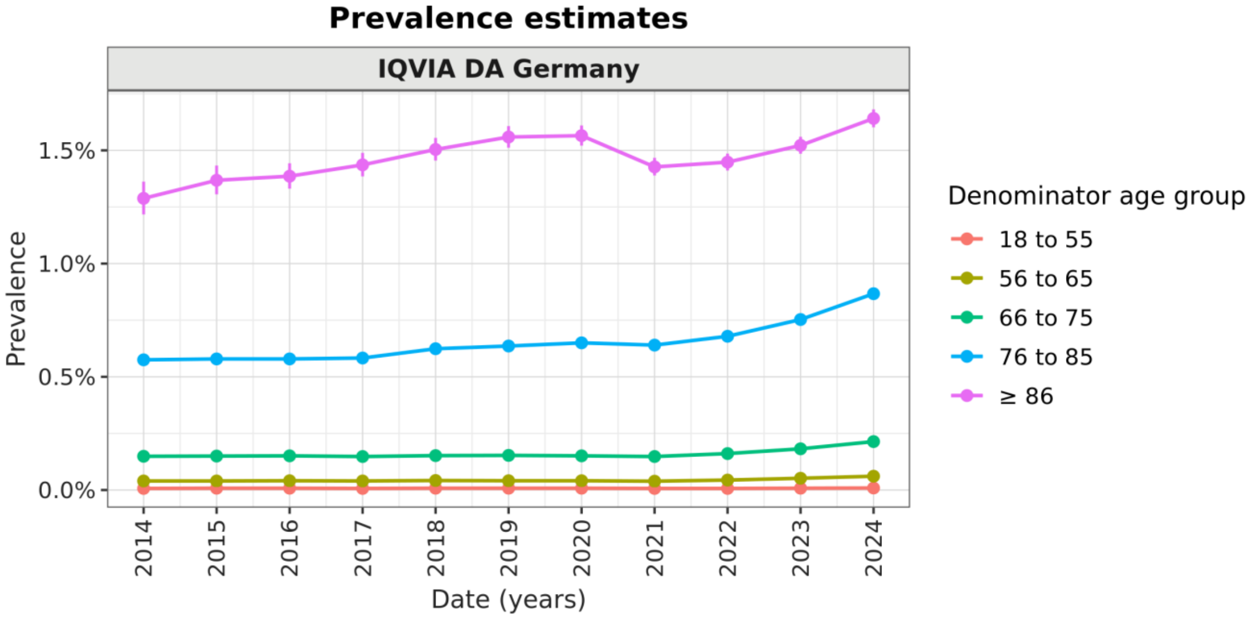

Figure 21. Prevalence (95% CIs) of AD in IQVIA DA Germany, stratified by age categories, 2014–2024.

IQVIA DA Germany=IQVIA Disease Analyzer Germany

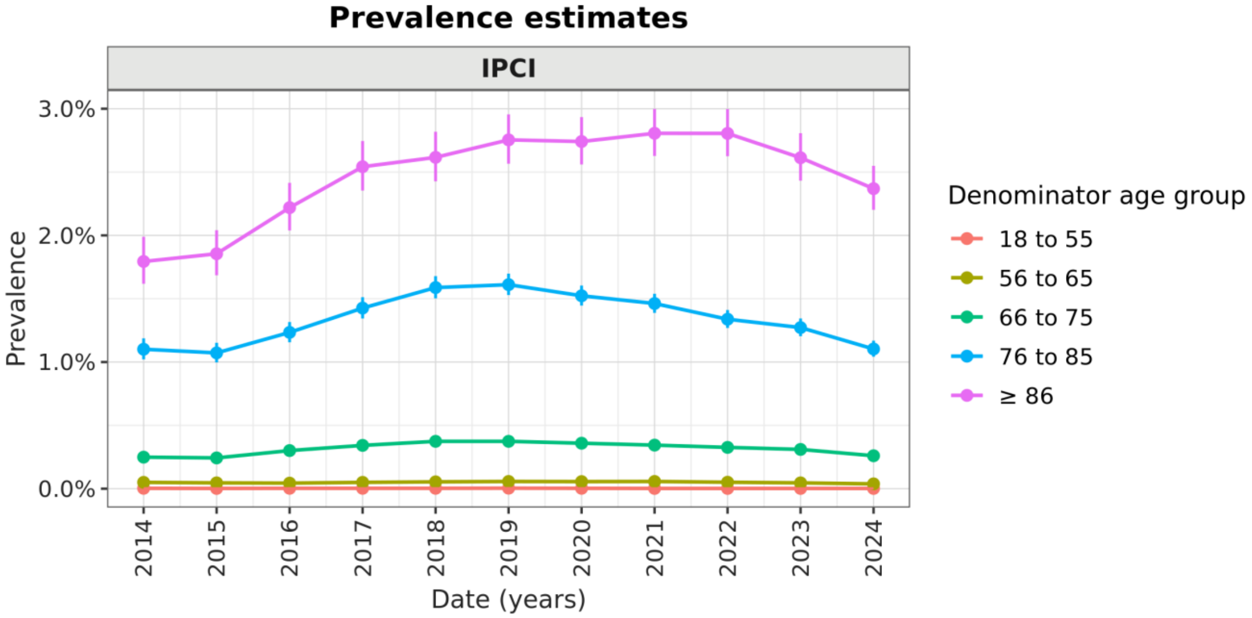

Figure 22. Prevalence (95% CIs) of AD in IPCI, stratified by age categories, 2014–2024.

IPCI=Integrated Primary Care Information

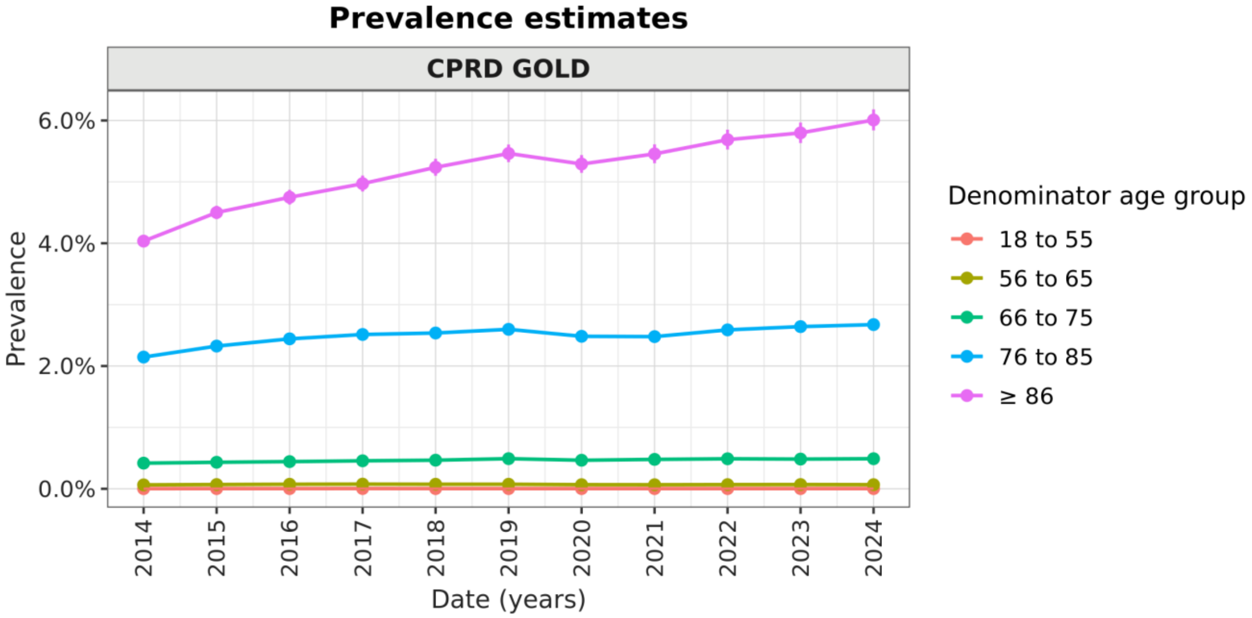

Figure 23. Prevalence (95% CIs) of AD in CPRD GOLD, stratified by age categories, 2014–2024.

CPRD=Clinical Practice Research Datalink

9.3.2. Objective 2 (Diagnostic procedures, individual-level, and clinical profile characterisation)

Large-scale characterisation of individuals with AD can be found in the ShinyApp at EUPAS1000000826. In short, the top three conditions were essential hypertension, dementia, and type 2 diabetes for NAJS; hypercholesterolemia, eczema, and essential hypertension, and cystitis in DK-DHR; essential hypertension, coronary arteriosclerosis, and hypokalaemia in InGef RDB; essential hypertension, nerve root disorder, and dementia in IQVIA DA Germany; essential hypertension, memory impairment, and urinary tract infectious disease in IPCI; and essential hypertension, amnesia, and eruption in CPRD GOLD. The top three medications were amoxicillin, diazepam, and dexamethasone for NAJS; acetaminophen, penicillin, and ibuprofen in DK-DHR; dipyrone, ibuprofen, and pantoprazole in InGef RDB; aspirin, dipyrone, and diclofenac in IQVIA DA Germany; omeprazole, aspirin, and simvastatin in IPCI; and amoxicillin, trimethoprim, and omeprazole in CPRD GOLD.

Pre-specified comorbidities

Table 5 reports the distribution of pre-specified comorbidities at any time prior to the diagnosis of AD. Across all data sources, the prevalence of comorbidities among individuals diagnosed with AD showed considerable variation, reflecting differences in population characteristics and healthcare settings. Down's syndrome was rare in all data sources, with the prevalence ranging from no counts in NAJS to approximately 0.2% in DK-DHR, IQVIA DA Germany, CPRD GOLD. Stroke was similarly uncommon, with most data sources reporting it in fewer than 0.3% of individuals with AD.

Cardiovascular comorbidities were common. The prevalence of atrial fibrillation showed variation, with the highest prevalence observed in InGef RDB (20.3%), followed by IPCI (10.2%) and CPRD GOLD (8.0%). Myocardial infarction was reported in 3.6–7.5% of individuals with AD across all data sources. The prevalence of heart failure ranged from 3.7% in CPRD GOLD to 21.9% in InGef RDB (**Table 5**).

Hypertension emerged as the most prevalent comorbidity across datasets, although its prevalence differed markedly. NAJS reported an exceptionally high prevalence of hypertension (83.9%), while the prevalence in other data sources ranged from 30.9% (CPRD GOLD) to 63.0% (InGef RDB). Diabetes was also common, with prevalence varying between 13.1% (CPRD GOLD) and 29.7% (NAJS) (**Table 5**).

The prevalence of lipid-related conditions displayed wide discrepancy between data sources. The prevalence of hypercholesterolemia ranged from 4.8% (NAJS) to 42.8% (DK-DHR). Hypertriglyceridemia was consistently rare, observed in less than 1.4% of individuals in all data sources (**Table 5**).

Table 5. Distribution of comorbidities at any time prior and up to diagnosis among individuals with Alzheimer's disease (number and %), by data source.

Comorbidities	NAJS (N=27,795)	DK-DHR (N=110,737)	InGef RDB (N=84,822)	IQVIA DA Germany (N=52,443)	IPCI (N=8,908)	CPRD GOLD (N=66,301)
Down's syndrome, N (%)	0 (0.00)	221 (0.20)	27 (0.03)	83 (0.16)	9 (0.10)	103 (0.16)
Stroke, N (%)	5 (0.02)	112 (0.10)	223 (0.26)	NA	<5	38 (0.06)
Atrial fibrillation, N (%)	1,563 (5.62)	4,535 (4.10)	17,216 (20.30)	3,038 (5.79)	909 (10.20)	5,304 (8.00)
Myocardial infarction, N (%)	1,307 (4.70)	8,249 (7.45)	5,942 (7.01)	2,344 (4.47)	439 (4.93)	2,372 (3.58)
Heart failure, N (%)	3,167 (11.39)	9,040 (8.16)	18,605 (21.93)	7,124 (13.58)	657 (7.38)	2,455 (3.70)
Hypertension, N (%)	23,324 (83.91)	54,374 (49.10)	53,468 (63.04)	25,600 (48.81)	3,282 (36.84)	20,456 (30.85)
Diabetes, N (%)	8,267 (29.74)	15,633 (14.12)	18,451 (21.75)	13,384 (25.52)	1,609 (18.06)	8,688 (13.10)
Hypercholesterolemia, N (%)	1,330 (4.79)	47,410 (42.81)	10,610 (12.51)	8,549 (16.30)	871 (9.78)	4,414 (6.66)
Hypertriglyceridemia, N (%)	194 (0.70)	81 (0.07)	343 (0.40)	702 (1.34)	14 (0.16)	47 (0.07)

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
NA=Not applicable (data not captured or poorly captured in the data sources)

Diagnostic procedures

The use of diagnostic procedures within 365 days prior to AD diagnosis was only available in InGef RDB and DK-DHR (**Table 6**). Brain MRI was performed in 18.0% (InGef RDB) and 13.9% (DK-DHR). Brain PET-F18 scans were uncommon, with DK-DHR showing the highest usage at 6.6%, followed by a small proportion in InGef RDB (0.2%).

Table 6. Distribution of diagnostic procedures within 365 days prior to diagnosis among individuals with Alzheimer's disease (number and %), by data source.

	Characteristics	NAJS (N=27,795)	DK-DHR (N=110,737)	InGef RDB (N=84,822)	IQVIA DA Germany (N=52,443)	IPCI (N=8,908)	CPRD GOLD (N=66,301)
Diagnostic procedure	Brain MRI, N (%)	NA	15,372 (13.88)	15,287 (18.02)	NA	NA	NA

	Characteristics	NAJS (N=27,795)	DK-DHR (N=110,737)	InGef RDB (N=84,822)	IQVIA DA Germany (N=52,443)	IPCI (N=8,908)	CPRD GOLD (N=66,301)
	Brain PET-F18, N (%)	NA	7,249 (6.55)	193 (0.23)	NA	NA	NA

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
MRI=Magnetic Resonance Imaging, PET-F18=Positron Emission Tomography with Fluorine-18, NA=Not applicable (data not captured or poorly captured in the data sources)

Clinical profile characteristics

Table 7 presents the distribution of clinical profile characteristics within 365 days prior to diagnosis among individuals with AD across six data sources. The table covers prespecified clinical features, use of AD-specific medications, and treatments commonly used for comorbid conditions.

Across the data sources, recording of caregiver support varied notably between sources: the prevalence was approximately 21.0% in InGef RDB and NAJS, 2.7% in IQVIA DA Germany, <1.0% in DK-DHR and CPRD GOLD, and not captured in IPCI. Alterations in activities of daily living were most frequently captured in NAJS (7.8%) and CPRD GOLD (4.1%), whereas proportions were low in DK-DHR, IQVIA DA Germany, and this information not captured in IPCI. MCI prior to AD diagnosis ranged from 1.9% (CPRD GOLD) to 4.4% (NAJS) and was not captured in IPCI (**Table 7**).

Use of AD medications also varied widely. Memantine was most frequently prescribed in NAJS (31.2%) and IQVIA DA Germany (8.9%), while donepezil was used more in DK-DHR (23.7%) and NAJS (17.6%). Rivastigmine and galantamine were prescribed at lower rates across all data sources, ranging between 0.4% and 6.0% for rivastigmine and 0.0–3.6% for galantamine (**Table 7**).

Medications related to comorbid conditions demonstrated substantial heterogeneity. Antiplatelet use ranged from 1.4% (NAJS) to 38.3% (DK-DHR). The prevalence of anticoagulant use was between 5.6% (IPCI) and 17.4% (InGef RDB). The prevalence of treatments for diabetes also showed wide variability, with InGef RDB reporting the highest proportions for both oral glucose-lowering drugs (7.8%) and insulin (8.9%). Antihypertensive medication use exceeded 40.0% in all datasets except IQVIA DA Germany (25.9%). Lipid-lowering drugs were most frequently prescribed in CPRD GOLD (41.8%) and InGef RDB (36.5%). Finally, treatments for heart failure were highly prevalent particularly in InGef RDB (77.0%) and NAJS (68.4%) (**Table 7**).

Table 7. Distribution of clinical profile characteristics within 365 days prior to diagnosis among individuals with Alzheimer's disease (number and %), by data source.

	Characteristics	NAJS (N=27,795)	DK-DHR (N=110,737)	InGef RDB (N=84,822)	IQVIA DA Germany (N=52,443)	IPCI (N=8,908)	CPRD GOLD (N=66,301)
Prespecified clinical features	Caregiver support, N (%)	5,714 (20.56)	683 (0.62)	17,922 (21.13)	1,403 (2.68)	NA	474 (0.71)
	Alterations in activities of daily living, N (%)	2,166 (7.79)	<5	NA	<5	NA	2,705 (4.08)
	MCI before AD diagnosis, N (%)	1,232 (4.43)	3,077 (2.78)	2,718 (3.20)	1,911 (3.64)	NA	1,263 (1.90)

Characteristics		NAJS (N=27,795)	DK-DHR (N=110,737)	InGef RDB (N=84,822)	IQVIA DA Germany (N=52,443)	IPCI (N=8,908)	CPRD GOLD (N=66,301)
Alzheimer's disease drugs	Memantine, N (%)	8,666 (31.18)	7,190 (6.49)	5,769 (6.80)	4,676 (8.92)	38 (0.43)	2,297 (3.46)
	Donepezil, N (%)	4,877 (17.55)	26,228 (23.68)	6,050 (7.13)	5,548 (10.58)	<5	7,246 (10.93)
	Rivastigmine, N (%)	122 (0.44)	5,659 (5.11)	3,526 (4.16)	3,165 (6.04)	281 (3.15)	669 (1.01)
	Galantamine, N (%)	0 (0.00)	1,584 (1.43)	1,325 (1.56)	1,862 (3.55)	211 (2.37)	1,168 (1.76)
Medications for comorbidities	Antiplatelets, N (%)	389 (1.40)	42,403 (38.29)	17,757 (20.93)	7,245 (13.81)	2,017 (22.64)	22,724 (34.27)
	Anticoagulants/antithrombotics, N (%)	3,731 (13.42)	14,958 (13.51)	14,772 (17.42)	3,726 (7.10)	499 (5.60)	6,198 (9.35)
	Oral glucose-lowering drugs, N (%)	1,984 (7.14)	3,156 (2.85)	6,579 (7.76)	1,846 (3.52)	369 (4.14)	3,051 (4.60)
	Insulin, N (%)	1,372 (4.94)	3,759 (3.39)	7,557 (8.91)	2,104 (4.01)	406 (4.56)	1,442 (2.17)
	Antihypertensives, N (%)	15,038 (54.10)	61,348 (55.40)	47,709 (56.25)	13,567 (25.87)	4,570 (51.30)	27,586 (41.61)
	Antiarrhythmics/rhythm control drugs, N (%)	909 (3.27)	971 (0.88)	1,303 (1.54)	423 (0.81)	362 (4.06)	718 (1.08)
	Lipid lowering drugs, N (%)	7,773 (27.97)	39,114 (35.32)	30,995 (36.54)	8,765 (16.71)	3,079 (34.56)	27,694 (41.77)
	Heart failure treatment, N (%)	18,997 (68.35)	59,859 (54.05)	65,327 (77.02)	18,435 (35.15)	4,890 (54.89)	29,929 (45.14)

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
NA=Not applicable (data not captured or poorly captured in the data sources)

Table 8 presents the distribution of prior cognitive diagnoses of MCI among individuals with AD, along with the duration between MCI diagnosis and subsequent AD diagnosis, overall and stratified by sex and age. The proportion of individuals with a recorded MCI diagnosis varied across data sources: 5.5% in NAJS, 4.2% in DK-DHR, 5.3% in InGef RDB, 8.3% in IQVIA DA Germany, and 4.4% in CPRD GOLD.

Across data sources in which MCI was recorded, the time from MCI to AD diagnosis showed wide variation. Median durations ranged from 351 days (NAJS) to 612 days (IQVIA DA Germany), with IQVIA DA Germany also showing the longest maximum duration (up to 9,038 days). Mean durations were similarly variable, from 583 days in NAJS to 896 days in IQVIA DA Germany (**Table 8**).

The median progression duration from MCI to AD was longer in females than males in NAJS (362 days for females and 311 days for males) and InGef RDB (405 days for females and 332 days for males), but shorter in DK-DHR (508 days for females and 558 days form males), IQVIA DA Germany (603 days for females and 625 days for males) and CPRD GOLD (448 days for females and 498 days for males) (**Table 8**).

Age-stratified results showed a clear increase in time to AD diagnosis with advancing age. In the youngest group (18–55 years), median durations ranged from 102 days (InGef RDB) to 760 days (CPRD GOLD). In the age groups of 56–65 years and 66–75 years, medians rose to 176–438 and 272–525 days, respectively. The oldest group (≥86 years) had the longest progression times, ranging from 401 days (NAJS) to 821 days (IQVIA DA Germany) (**Table 8**).

Table 8. Distribution of prior cognitive diagnosis (MCI) at any time prior and up to diagnosis among individuals with Alzheimer's disease (number and %), by data source, sex, and age category.

	Characteristics	**NAJS**	**DK-DHR**	**InGef RDB**	**IQVIA DA Germany**	**IPCI**	**CPRD GOLD**
Overall	MCI, N (%)	1,517 (5.46)	4,656 (4.20)	4,482 (5.28)	4,345 (8.29)	NA	2,908 (4.39)
	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–3,547	0–9,216	0–3,367	0–9,038	NA	0–5,781
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	351 (98–840)	532 (224–1,097)	370 (79–915)	612 (217–1,228)	NA	469 (222–922)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	583.29 (642.88)	831.14 (936.80)	599.61 (651.97)	895.61 (955.31)	NA	683.66 (670.90)
Sex							
Female	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–3,352	0–9,216	0–3,241	0–7,996	NA	0–5,503
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	362 (97–830)	508 (218–1,068)	405 (85–946)	603 (204–1,239)	NA	448 (212–887)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	583.72 (635.45)	814.58 (922.65)	619.00 (653.79)	897.36 (960.49)	NA	662.46 (666.47)
Male	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–3,547	0–8,897	0–3,367	0–9,038	NA	0–5,781
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	311 (99–865)	558 (231–1,134)	332 (73–858)	625 (233–1,213)	NA	498 (236–982)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	582.45 (657.66)	855.63 (957.08)	575.64 (649.09)	892.90 (948.80)	NA	715.87 (676.61)
Age category							
18–55	Duration in days from MCI occurrence to AD diagnosis (min – max)	42–1,892	0– 4,948	0–2,032	0–3,330	NA	0–1,825
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	256 (127–488)	302 (73–673)	102 (0–372)	305 (12–960)	NA	760 (157–1,412)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	453.77 (505.06)	560.19 (899.82)	298.68 (437.93)	580.30 (760.50)	NA	821.67 (774.55)

	Characteristics	NAJS	DK-DHR	InGef RDB	IQVIA DA Germany	IPCI	CPRD GOLD
56–65	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–2,387	0–4,724	0–3,137	0–4,652	NA	0–2,657
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	362 (122 – 831)	409 (116–782)	176 (42–625)	438 (140–903)	NA	362 (206–648)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	542.22 (547.73)	591.79 (687.28)	445.57 (593.63)	701.19 (826.19)	NA	526.69 (535.18)
66–75	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–2,899	0–6,929	0–3,225	0–5,337	NA	0–3,944
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	336 (95–802)	482 (189–1,001)	272 (62–740)	524 (182–1,039)	NA	452 (221–865)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	564.99 (628.48)	739.08 (832.89)	508.91 (597.83)	763.05 (823.47)	NA	638.61 (583.56)
76–85	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–3,547	0–9,216	0–3,367	0–7,996	NA	0–4,257
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	356 (97–868)	554 (245–1,118)	416 (94–927)	616 (243–1,231)	NA	478 (223–930)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	601.41 (662.78)	860.68 (967.77)	614.99 (646.13)	888.97 (911.97)	NA	678.09 (641.25)
≥86	Duration in days from MCI occurrence to AD diagnosis (min – max)	0–3,352	0–6,635	0–3,241	0–9,038	NA	0–5,781
	Duration in days from MCI occurrence to AD diagnosis (median (IQR))	400 (93–811)	612, (262–1,322)	610 (150–1,176)	821 (314–1,572)	NA	476 (221–984)
	Duration in days from MCI occurrence to AD diagnosis (mean (SD))	628.00 (707.69)	963.37 (1,026.95)	774.19 (721.29)	1,175.29 (1,219.78)	NA	742.67 (790.88)

NAJS=National Public Health Information System, DK-DHR=Danish Data Health Registries, InGef RDB=InGef Research Database, IQVIA DA Germany=IQVIA Disease Analyzer Germany, IPCI=Integrated Primary Care Information, CPRD=Clinical Practice Research Datalink
AD=Alzheimer's disease, IQR= interquartile range, MCI=Mild Cognitive Impairment, SD=Standard Deviation, NA=Not applicable (data not captured or poorly captured in the data sources)

9.4. Other analysis

None.

10. DISCUSSION

10.1. Key results

Across six European data sources, individuals with AD were consistently older (median 79–82 years) than the general adult population (median 34–51 years) and predominantly female (56–66%).

The overall incidence rate of AD in the adult population (aged ≥18 years) ranged from 53.85 (IQVIA DA Germany) to 137.33 (InGef RDB) per 100,000 PY and was higher among females than males and increasing with age. There was no clear trend (increase or decrease) in the incidence of AD over the years across the data sources. The overall prevalence of AD ranged from 0.16% (IQVIA DA Germany) to 1.66% (InGef RDB) and was also higher among females than males and increased with age. The prevalence of AD showed a higher increase over the years for NAJS and InGef RDB and was more stable for the remaining data sources.

Hypertension (30.9–83.9%), diabetes (13.1–29.7%), and cardiovascular diseases (atrial fibrillation ranging from 4.1–20.3% and heart failure 3.7–21.9%) were common comorbidities among individuals with AD. Use of brain MRI ranged from 13.9–18.0%, and PET-F18 scans were rarely used, except in DK-DHR (6.6%). Caregiver support was reported in 0.7–21.1% of cases, daily activity limitations 0.0–7.8%, and prior MCI diagnoses in 1.9–4.4%. Memantine (0.4–31.2%) and donepezil (7.1–23.7%) were more commonly used AD drugs, while rivastigmine and galantamine use remained low (0.0–6.0%). Treatments for comorbidities included antiplatelets use (1.4–38.3%), anticoagulants use (5.6–17.4%), antihypertensives (>40.0% in most sources), and heart-failure therapies (35.2–77.0%). Among individuals with prior MCI, time to AD diagnosis varied considerably, with median intervals from approximately 350 to over 600 days, in which the time increased with age.

Incidence and prevalence of AD in the general adult population

The overall incidence of AD ranged from 41.51 per 100,000 PY for IQVIA DA Germany to 198.41 per 100,000 PY for InGef RDB. The incidence was higher among females than males and ranged from 39.00 to 221.12 per 100,000 PY for females and from 33.91 to 175.42 per 100,000 PY for males. A higher incidence among females compared to males was also observed among most age groups. The incidence of early-onset AD ranged from 2.35 to 12.24 per 100,000 PY, while this was much greater for late-onset AD, ranging from 124.79 to 1,146.32 per 100,000 PY. When stratifying the results by age, all data sources showed an increase of the incidence of AD with increasing age, with a significantly higher incidence among the two oldest age groups (76–85 and 86 years and over), than the younger age groups (18–55, 56–65, and 66–75 years).

The overall prevalence of AD ranged from 0.11% for IQVIA DA Germany to 0.69% for DK-DHR. The prevalence of AD was higher among females than males among all data sources and ranged from 0.11% to 0.83% for females and from 0.10% to 0.56% for males. A higher prevalence among females compared to males was generally also observed among each age group. The prevalence of early-onset AD was found to be 0.00% among all data sources, while for late-onset AD the prevalence ranged from 0.36% in IQVIA DA Germany to 3.23% in InGef RDB. When stratifying age into more categories, a higher prevalence of AD was observed among the two oldest age groups (76–85 and 86 years and over), than the younger age groups (18–55, 56–65, and 66–75 years).

Diagnostic procedures and clinical profile of individuals who are diagnosed with incident AD

Comorbidity patterns showed substantial variability. Cardiovascular conditions were common, with atrial fibrillation up to 20.3% and heart failure up to 21.9% in some data sources. Hypertension was the most prevalent comorbidity, ranging from 30.8–83.9%. Diabetes affected 13.1–29.7% of individuals with AD, while Down's syndrome and stroke were rare. Lipid disorders showed strong variation, with hypercholesterolemia ranging from 4.8–42.8%.

Diagnostic procedures before AD diagnosis differed markedly across settings. Brain MRI use ranged from 0.07% to 18.0%, with the highest use in InGef RDB and DK-DHR. PET-F18 imaging was rarely used, recorded mainly in DK-DHR (6.6%), with minimal or no use in other data sources.

The clinical profile of individuals with AD varied across data sources. Caregiver support ranged from 1–21.1%, with the highest levels in NAJS and InGef RDB. Daily activity limitations were most common in NAJS

(7.8%) and CPRD GOLD (4.1%), with minimal recording elsewhere. Prior MCI diagnosis ranged from 1.9–4.4% where captured, with none in DK-DHR or IPCI. Use of AD medications differed notably: memantine ranged from 0.4–31.2%, donepezil from <5 cases–23.7%, and rivastigmine and galantamine remained low (0.4–6.0% and 0.0–3.6%, respectively). Comorbidity treatments showed wide variation, with antiplatelet use at 1.4–38.3%, anticoagulants at 5.6–17.4%, and diabetes medications 2.2–8.9%. Antihypertensive use exceeded 40.0% in all data sources except IQVIA DA Germany (25.9%). Lipid-lowering therapy ranged from 16.7–41.8%, and heart-failure treatment was particularly common in InGef RDB (77.0%) and NAJS (68.4%).

Among individuals with prior MCI, the time from MCI to AD diagnosis differed substantially. Median progression ranged from approximately 350 days (NAJS) to over 600 days (IQVIA DA Germany). No difference in duration was observed consistently between females and males across data sources, but older age groups had longer intervals between MCI and AD, with medians ranging from 400–821 days in individuals aged ≥86 years compared to 102–760 among individuals aged 18–55 years.

10.2. Strengths and limitations of the research methods

As the majority of publications focus mainly on dementia in general, limited studies are available on AD specifically, which makes the current study of added value. The study was performed using routinely collected health care data and so, data quality issues must be considered when interpreting the results. Electronic health records have certain inherent limitations because they were collected for clinical purpose rather than primarily for research use. In general, data on genetics (e.g., APOE4, APP, PSEN1, PSEN2), AD-specific biomarkers (amyloid-β, tau), disease staging, and functional or lifestyle factors were either unavailable or inconsistently captured across the selected data sources, restricting the depth of characterisation of individuals. Regarding the variables included in this study, information on ADL and diagnostic procedures was sparse or not available in the data sources. Variability across data sources in coding practices, availability of radiology, and completeness of comorbidity information may introduce heterogeneity and potential misclassification, particularly when distinguishing AD from other dementias. This study was based on a phenotype definition which only included AD disease (code ID: 378419). Therefore, early AD cases coded as “unspecified dementia” may have been missed. Furthermore, while OMOP provides mapping to established vocabularies such as SNOMED, inaccuracies or gaps in these mappings can occur, impacting the accuracy and completeness of data in different data sources. This was observed for data sources of which the vocabularies are based on International Classification of Diseases, Tenth Revision: in NAJS, DK-DHR, InGef RDB, and IQVIA DA Germany, all available G30 source codes were mapped to the AD concept code. As a result, the less specific codes for “Other Alzheimer disease” (G30.8) and “Alzheimer disease, unspecified” (G30.9) were included in the AD cohort definition used in this study. While these codes have some etiological overlap with AD, they represent broader diagnostic categories that may not be specific to AD. However, within the OMOP CDM infrastructure available in ATHENA,[10] there are no alternative mappings available. Given this technical constraint, it was not possible to exclude the specific codes from the aforementioned data sources.

Another limitation is that in the InGef RDB claims data source, there is potential misclassification of the dates of diagnosis for conditions diagnosed outside of a hospital setting. Unlike inpatient diagnoses, which are associated with an exact diagnosis date and mapped directly to standard concepts, outpatient diagnoses are recorded on a quarterly basis without a specific date. All outpatient diagnoses within a given quarter are documented as occurring on the last day of that period. For example, if an individual is first diagnosed with AD during an outpatient visit on 12 May, the diagnosis will be recorded as occurring on 30 June, the last day of the second quarter. This misclassification can lead to a time lag of up to three months between the date of actual diagnosis and the recorded date in the data source. This can, for example, impact results where the time is calculated from MCI occurrence to AD diagnosis, in which the time was not accurately calculated.

In addition, there might be a misclassification of prevalent AD cases as incident AD cases in the first years of the study period for InGef RDB. This is due to the fact that data was available in InGeF RDB starting from 01/01/2015 with no previous information available on, for example, demographics, diagnoses, and treatments. This means that individuals could have been accounted as incident AD cases (due to first occurrence in the data source), while in fact they were diagnosed with AD before and thus should have been accounted as prevalent AD cases. This would explain why in the first years of the study period, a sharp decrease was observed in the incidence of AD and a sharp increase of the prevalence of AD. This would mean that the more stable incidence and prevalence estimates observed over the last years of the study period more closely represent the accurate estimates of these values, as patients have longer pre-observation periods in which prior AD diagnoses can be identified.

10.3. Interpretation

Across six European real-world data sources, the epidemiology of AD showed consistent patterns: incidence and prevalence increased sharply with age, with the highest rates observed among individuals aged 76 years and older. Both measures were higher in females than males, and although absolute rates varied across data sources, the age-related trends were uniform. Individuals with AD commonly had cardiovascular and metabolic comorbidities, and diagnostic practices and treatment patterns differed substantially between data sources. Overall, these findings align with established evidence on the strong age dependence and clinical complexity of AD, while highlighting important cross-country variability in detection and management. Our findings also underscore the public-health implications of demographic aging: as life expectancy continues to increase, the number of individuals entering the highest-risk age groups will rise, contributing to a growing burden of AD in Europe, a pattern mirrored in global projections.[11]

The observed increase in AD incidence with advancing age across all data sources aligns closely with established epidemiological evidence. Numerous longitudinal and population-based studies have shown that AD incidence rises sharply with age, often accelerating after the age of 75 years. For example, a community-based cohort study showed nearly exponential increases in incidence from the late 60s through the age of 90 years and above, supporting the steep age-related gradient observed in our datasets.[12] Similar patterns were reported in large European pooled analyses, where incidence increased from approximately 3–4 per 1,000 PYs in individuals aged 65–74 years to more than 35 per 1,000 PYs among those aged 85 and older.[13] A meta-analysis on the incidence and prevalence of AD in Europe showed an incidence of 1,108 (1,030–1,189) per 100,000 PY and a prevalence of 5.05% (4.73–5.39%).[14] This meta-analysis was based on publications including populations aged 55 years and older, while our study included populations aged 18 years and older. This study also showed higher incidence, and prevalence estimates for females compared to males and with increasing age,[14] which is in line with our study.

Although the absolute incidence levels varied across the participating data sources, likely reflecting differences in healthcare systems, diagnostic practices, and data capture, the consistent age-related trends suggest a robust underlying epidemiological signal. Prior literature has highlighted such variability between populations and registries, reinforcing the need for harmonised definitions and cross-country comparisons when interpreting AD epidemiology. The heterogeneity of the data sources included in this study can play a substantial role on the incidence and prevalence estimates of AD. The included data sources differ in their reliance on primary care records, ambulatory secondary (specialist) care, and hospital inpatient data, each of which captures AD at different stages of the clinical care pathway. Primary care systems (e.g., IPCI and IQVIA DA Germany) would mainly record milder and earlier AD disease presentations (e.g., initial cognitive symptoms) but may underreport the more severe or complex cases if diagnoses are confirmed in specialty or hospital settings and not consistently reported back to the GP. In addition, cognitive symptoms may be documented under nonspecific codes, leading to variability in incidence estimates. On the other hand, specialist and hospital systems (e.g., DK-DHR) would mainly capture more advanced (moderate-to-severe cases) and confirmed AD cases but could miss initial AD presentations. Furthermore, hybrid systems (e.g., InGef RDB and NAJS) that integrate multiple care settings (e.g., primary and secondary care) capture a broader AD case spectrum.

While InGef RDB and IQVIA DA Germany are data sources from the same country (Germany), the difference in the nature of the data source most likely led to the differences in the incidence and prevalence estimates. The lower estimates in IQVIA DA Germany compared to the data from InGef RDB could be attributed to the outpatient coverage of IQVIA DA Germany which primarily captures records from GPs and primary care specialists, while InGef RDB includes both inpatient and outpatient data. In addition, the records between health care providers are not linked in IQVIA DA Germany. A lack of feedback of the diagnosis of AD from a specialist setting to the GP could also explain the generally lower estimates observed in the sources of GP data (i.e., IPCI and IQVIA DA Germany).

With respect to the characterisation of individuals with AD, the diagnostic procedures (brain MRI and brain PET-F18) and prespecified clinical features (caregiver support, alterations in activities of daily living, and MCI before AD diagnosis) were in general not available in the data sources or information was very limited or only captured for specific populations, hence the low counts. Furthermore, hypercholesterolemia and hypertriglyceridemia diagnosis were in this study based on conditions (SNOMED vocabulary), while ideally also measurements (for example blood tests) should be used. Therefore, these variables might not be truly representative and should be interpreted with caution. NAJS reported an exceptionally high prevalence of hypertension (83.9%), however, this represents both suspected and confirmed cases. Hypertension is based on solely ICD-10 diagnosis, without antihypertensive therapy. Having only a diagnosis of hypertension without antihypertensive therapy should not be considered as a hypertension case in NAJS, as the diagnosis for hypertension is also used in primary health care data as a working diagnosis or can be registered for any prior visit to a primary care for an individual who had suspected hypertension. For example, a person who visits the GP with high blood pressure at one point would have a working diagnosis of hypertension. However, if it appears that this person did not have high blood pressure afterwards and did not receive any therapy for it, it would still be evidenced in the primary healthcare data as a visit with a diagnosis of hypertension even though it is not a confirmed diagnosis and no therapy was prescribed. Furthermore, the prevalence of hypercholesterolemia was relatively low in NAJS, which is most likely due to its registration as "Disorders of lipoprotein metabolism and other lipidemias" (ICD-10 code E78), which was not used in this study. The use of lipid lowering drugs (see Table 7) is a better indicator of the population which has high cholesterol or triglycerides.

Importantly, the time interval between MCI and AD diagnosis was long and varied across data sources, ranging from approximately one to nearly two years depending on the data source. This variation may reflect delays in clinical recognition, differences in follow-up intensity, or differing thresholds for diagnosing AD. The median time (in days) from MCI to AD diagnosis was observed to be increasing with age (however, this increase is less clear when taking into account the IQR in NAJS and CPRD GOLD), consistent with previous studies that suggested the progression rates from MCI to AD are influenced by age, comorbidities, and healthcare utilisation patterns.[15] The increasing time from MCI to AD diagnosis with age likely reflects diagnostic and healthcare system factors rather than true differences in disease progression. Many comorbidities, including hypertension and hyperlipidaemia, share risk factors (such as aging) with AD and may influence AD onset. Furthermore, detection bias is likely, as individuals experiencing cognitive decline often undergo more rigorous medical evaluations, increasing the likelihood of identifying comorbidities.[15] These various underlying pathologies can challenge the diagnosis of AD.

10.4. Generalisability

Due to feasibility and timelines, only six primary care/secondary care data sources from Croatia, Denmark, Germany, The Netherlands, and the United Kingdom could be included, which operate under blanket IRB approval or can obtain approval within a month. This limited the generalisability of the findings to these healthcare systems and populations. Follow-up studies may include additional data sources to broaden the generalisability of the findings and inclusion of data sources with more detailed information on demographic characteristics, diagnostic procedures, and the clinical profile of individuals with newly diagnosed AD.

11. CONCLUSION

This multi-country analysis provides a comprehensive overview of AD epidemiology and clinical characteristics across diverse European healthcare settings. Despite variations in absolute rates and diagnostic practices, the findings consistently demonstrate that AD is predominantly a disease of older age, with incidence and prevalence rising steeply in the oldest age groups and higher rates observed among females. Individuals with AD frequently present with substantial cardiovascular and metabolic comorbidity, underscoring the need for integrated care approaches. Differences in diagnostic pathways, imaging use, and treatment patterns highlight opportunities for harmonising clinical practice and improving early detection. There is a need to improve the capture of relevant information in healthcare systems, such as genetics (e.g., APOE4, APP, PSEN1, PSEN2), AD-specific biomarkers (amyloid-β, tau), disease staging, and functional or lifestyle factors, to better characterise individuals with AD and to allow follow-up of the progression of the disease. Future studies can be conducted that would ideally include these relevant characteristics and that include hybrid data sources (i.e. data linked from multiple healthcare settings) to capture a broader AD case spectrum. Overall, the results reinforce the growing public-health burden of AD in ageing populations and emphasise the importance of coordinated strategies for prevention, diagnosis, and long-term management across Europe.

12. REFERENCES

1.Breijyeh Z KR. Comprehensive Review on Alzheimer's Disease: Causes and Treatment. Molecules. 2020;25(24):5789.

2.Nichols E SJ, Vollset SE, Fukutaki K, Chalek J, Abd-Allah F, et al. . stimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. The Lancet Public Health. 2022;7(2):e105-e25.

3.Liu W DW, Gong X, Ou J, Yu S, Chen S. Global burden of Alzheimer’s disease and other dementias in adults aged 65 years and over, and health inequality related to SDI, 1990–2021: analysis of data from GBD 2021. BMC Public Health. 2025;25(1):1256.

4.Wolters FJ CL, Waziry R, Anderson R, Berr C, Beiser A, et al. . Twenty-seven-year time trends in dementia incidence in Europe and the United States: The Alzheimer Cohorts Consortium. Neurology. 2020;95(5):e519-e31.

5.Yu DT LR, Sun JR, Rong XW, Guo XG, Zhu GD. Global mortality, prevalence and disability-adjusted life years of Alzheimer's disease and other dementias in adults aged 60 years or older, and the impact of the COVID-19 pandemic: a comprehensive analysis for the global burden of disease 2021. BMC Psychiatry. 2025;25(1):503.

6.Dernie F CG, Robinson A, Bezer J, Mercade-Besora N, Griffier R, et al. Standardised and Reproducible Phenotyping Using Distributed Analytics and Tools in the Data Analysis and Real World Interrogation Network (DARWIN EU). Pharmacoepidemiol Drug Saf. 2024;33(11):e70042.

7.Raventós B CM, Du M, Guo Y, Black A, Inberg G, et al. IncidencePrevalence: An R package to calculate population-level incidence rates and prevalence using the OMOP common data model. Pharmacoepidemiology and Drug Safety. 2024;33(1):e5717.

8.Catala M GY, Lopez-Guell K, Burn E, Mercade-Besora N, Alcalde M, et al. CohortCharacteristics: Summarise and Visualise Characteristics of Patients in the OMOP CDM. R package version 1.0.1. 2025 Available from: https://darwin-eu.github.io/CohortCharacteristics/.

9.Català M GY, Du M, Lopez-Guell K, Burn E, Mercade-Besora N. Patientprofiles: Identify Characteristics of Patients in the OMOP Common Data Model. R package version 1.4.3 2025. Available from: https://darwin-eu.github.io/patientprofiles/.

10. Observational Health Data Sciences and Informatics (OHDSI). Athena - OHDSI vocabularies repository. Available from: https://athena.ohdsi.org/search-terms/terms.

11. 2025 Alzheimer's disease facts and figures: Alzheimers Dement. 2025 Apr 29;21(4):e70235. doi: 10.1002/alz.70235. eCollection 2025 Apr.

12. Qiu C, Kivipelto M, von Strauss E. Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. Dialogues Clin Neurosci. 2009;11(2):111-28.

13. Tahami Monfared AA, Byrnes MJ, White LA, Zhang Q. Alzheimer’s Disease: Epidemiology and Clinical Progression. Neurology and Therapy. 2022;11(2):553-69.

14. Niu H, Álvarez-Álvarez I, Guillén-Grima F, Aguinaga-Ontoso I. Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. Neurologia. 2017;32(8):523-32.

15. Rogers B. Unraveling temporal patterns of diagnostic markers and comorbidities in Alzheimer's disease: Insights from large-scale data. Alzheimer's & Dementia. 2025;21(3):e14564.

13. ANNEXES

ANNEX I. Description of data sources

Croatia, National Public Health Information System (NAJS)

#	Section	Description
1	Data source identification and country	NAJS (Croatian National Public Health Information System) Croatia
2	Data partner information section	Croatian Institute of Public Health Department of Data Science and Analytics
3	Coverage and timespan	Data collection since: 1998 Extent: Nation-wide. Geographic coverage covers whole Croatia, with various levels of resolution for different registries. Current estimates for the population in Croatia will be available at: https://podaci.dzs.hr/hr/podaci/stanovnistvo/procjena-stanovnistva/ for each year. The total and active person count in the NAJS data is larger than the current population of Croatia. This explained by: a) the person table included deceased and all previously insured people and b) there is no information about insurance ending, c) healthcare is also used by people with dual citizenship from neighbouring countries It is known that a lot of people emigrated (300k-400k) and weren't included in the last population census but still are in the NAJS database. There is also an influx of immigrant workers that are insured and registered but weren't included in the census.
4	Healthcare setting / type of data	Primary care – General Practioner, and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. Primary care – gps, and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. For both inpatient and outpatient setting diagnoses, medication, procedures, and measurements are captured. The year of availability of information depends on the setting • 2014-2025 for biochemical lab tests in primary care from EHR patients records (measurements with results) • 2015-2025 for primary care data from EHR patient records (conditions, procedures, and drug prescriptions) • 2015- 2024 for inpatient hospital data from EHR administrative records (conditions, procedures, measurements without results and drug administrations) • 2016-2025 for health risk assessment data entered by GPs (measurements with results - height, weight…) • 2016-2022 for secondary conciliatory care data from EHR administrative records (conditions, procedures, measurements without results and drug administrations) • 2016-2022 for emergency care data from EHR patient records (conditions) • 2017-2025 for hospital records from registry data (conditions and procedures) • 2020-2025 for vaccination data from EHR patient records
5	Data collection process	Inpatient hospital billing systems, and Other. Data is entered by clinicians at healthcare contact, then combined by CIPH into the NAJS database and integrated with registries for public health purposes.
6	General representativeness	The data is collected from the evidence of public health records collected for public health purposes, as the majority of health care in Croatia is public and under single health insurance provider. Personal details are collected to a better extent for insured individuals compared to uninsured patients, who are excluded in the ETL process.
7	Data content /source coding	Medication prescriptions are recorded with ATC codes with an additional 3 digit code denoting the package. Diagnoses with ICD10 codes (Australian modification). Procedures with local source codes. Lab results with local source codes.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. Records from 2017 include insured patients with reliable IDs. Uninsured patients do not have reliable IDs. For example, if a patient changed her status from insured to uninsured, or vice versa,

#	Section	Description
		she could be counted several times, as could tracking records from before 2017 and after. By using the unique personal identifier for Croatian citizens, it can be checked and verified.
9	Quality control (data source specific)	There is a network of registry personnel (leaders, administrators, coders, sources) working on data coverage and other quality dimensions. An analytical team routinely checks for erroneous entries in hospital records, removing double entries, false dates, and overlapping stays. Entries without enough data or with obviously erroneous dates from primary care analysis are being excluded.
10	Linkage	The national death registry is updated yearly, with one year lag, but the fact of someone's death (just the date) is updated daily, without the cause of death or any other additional details. Primary care is updated weekly and hospital level care monthly. Specific registries are included in NAJS (e.g. diabetes registry), where inclusion criteria vary across these registries.
11	Vital status	NAJS is linked to the national death registry.
12	Limitations	Hospital data is available from 2017 onwards. This is often used as start of data collection, while laboratory and GP data is captured before that (since 2014 and 2015 respectively). Drug duration is often not available and set to 1 day for administration and 30 days for prescription. Hospital discharge summaries are currently not captured in NAJS. Hospital drug administration data is less reliable than prescription data from primary care, with some drugs (monoclonal antibodies / precision medicine drugs) that require additional approval not being recorded at all.
13	Main references	No main reference provided.
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111155 Website: https://www.hzjz.hr/nacionalni-javnozdravstveni-informacijski-sustav-najs/

Denmark, Danish Data Health Registries (DK-DHR)

#	Section	Description
1	Data source identification and country	DK-DHR (Danish Data Health Registries) Denmark
2	Data partner information section	Danish Medicines Agency (DKMA) Data Analytics Centre (DAC)
3	Coverage and timespan	Data collection since: 1995 Extent: Nation-wide. The data is representative of the entire Danish population.
4	Healthcare setting / type of data	Community pharmacists, and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. The following data elements are collected: diagnosis (including rare diseases and pregnancy data), hospital admissions, discharge and ICU data, Cause of death, Drug prescriptions, dispensing, vaccination and contraception, Procedures (surgical and non-surgical hospital), and Sociodemographic information (sex and age only).
5	Data collection process	Outpatient electronic health records, and Inpatient hospital electronic health records, and Registries, and Other. All causes of deaths, all retrieved drug prescriptions, all records of vaccinations, all hospital inpatient and outpatients contacts including disease diagnoses and hospital surgical and non-surgical procedures, histologically confirmed incident cancers, laboratory test results for the entire Danish population from 1/1/1995 onwards.
6	General representativeness	The data is representative of the entire Danish population. Healthcare is free in Denmark, so we do not expect any bias in data collection based on socio-economic status.

#	Section	Description
7	Data content /source coding	Diagnoses and causes of death are collected using the ICD-10 vocabulary. ATC and RxNorm are used for Drugs. SNOMED codes are used for Procedures.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. Patients have unique identifiers used to link datasets.
9	Quality control (data source specific)	The data we have received relating to nationwide Danish Health Data registries offer an opportunity for large-scale, population-based studies with several advantages 1) Their large size improves the precision of estimates and enables the study of rare exposures and outcomes with long-term latency, 2) Inclusion of nearly all individuals in the target population ensures that the data reflect routine clinical care and all clinical segments of the source population, 3) Data are collected independently of each research study, thus minimising certain types of bias, e.g., non-response, and the influence from attention to the research question on the diagnostic process. Before the source data is sent to us, the Danish Health Data Authority runs and does comprehensive checks of the registry table data validity of the variables, breaks in data, changes in variable coding, missingness, etc. We perform checks of missingness/completeness in relation to requested variables. In essence, we are receiving a dump of a mirror of the data that is controlled by the SDS. The documentation performed by SDS is available online, in Danish primarily https://www.esundhed.dk/Dokumentation (all variables), but also in English https://sundhedsdatastyrelsen.dk/da/english/health_data_and_registers/national_health_registers
10	Linkage	There is no linkage in this data source.
11	Vital status	The Cause of Death registry (DAR) is used, the cause of death is collected using ICD-10 codes.
12	Limitations	DK-DHR has the following limitations, which may be relevant confounders for certain complex Darwin EU studies: - We lack information on key socio-economic status (SES) factors, such as occupation, education, and income. These variables may be important for analysis in some studies. - We only have complete data on lifestyle factors (such as smoking status and weight) for pregnant women. - We have no information on patient contacts in primary care (visits to the GP). Consequently, the incidence of chronic diseases like Type 2 Diabetes (T2D) and asthma must be determined using drug prescriptions as a proxy. Stillborn children will not have any records in our CDM. This means that e.g. birth length of stillborns is not recorded.
13	Main references	Schmidt M, Schmidt SAJ,Adelborg K,Sundbøll J,Laugesen K,Ehrenstein V,Sørensen HT "The Danish health care system and epidemiological research: from health care contacts to database records." Clinical epidemiology (2019): 31372058
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111217 Website: https://sundhedsdatastyrelsen.dk/da/english/health_data_and_registers/healthdatadenmark

Germany, InGef Research Database (InGef RDB)

#	Section	Description
1	Data source identification and country	InGef RDB (InGef Research Database) Germany
2	Data partner information section	Institut für angewandte Gesundheitsforschung Berlin GmbH
3	Coverage and timespan	Data collection since: 2014 Extent: Nation-wide. The data source contains information from the statutory health insurances (SHI), which insure a total of about 89% (~73 million individuals) of the German population. Since the InGef RDC currently includes about ten million individuals, it covers about 13% of the total population

#	Section	Description
		insured in one of the German SHIs. The data in the database depicts all health care use which has been reimbursed by the SHI.
4	Healthcare setting / type of data	Primary care – General Practioner, and community pharmacists, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care, and claims data. The following data elements are presented in the OMOP CDM: demographic information, diagnoses, procedures, dispensing drugs and advanced therapy medicinal products, vaccinations, pregnancy data (via diagnoses and procedures) and contraception.
5	Data collection process	Insurance/administrative claims. The data in the database depicts all health care use which has been reimbursed by the SHI (statutory health insurances).
6	General representativeness	The RDB covers about 11% of the German population and is comparable to the German population in terms of the distribution of age and sex. Most health insurances that contribute to the RDB have nationwide coverage, meaning that the database covers all regions of Germany. Since almost all services covered by statutory health insurances are specified in national legislation, healthcare provision all over Germany is well represented in the RDB. Additionally, in Germany it is very common to stay with the same health insurance throughout life, which results in a good longitudinal coverage over the entire period of 10 years.
7	Data content /source coding	The coding in the research database complies with national classification and coding rules in Germany. Diagnoses are coded according to ICD-10-GM. Inpatient and outpatient surgeries or procedures are recorded as OPS codes (German classification of Operations and Procedures). The dispensing of drugs in pharmacies is recorded using the PZN (pharmaceutical registration number). For drugs that miss a PZN-to-RxNorm mapping, the ATC code is used instead. In some cases, dispensed drugs can be coded using OPS codes (e.g. in hospitals) or EBM codes (fee schedule for outpatient treatments).
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. In the German statutory health system, a person can only be enrolled in one health insurance at a time. However, if a person changes from one contributing insurer to another, a new ID number will be generated.
9	Quality control (data source specific)	The data transmitted by healthcare providers complies with the standardized requirements and formats of the Association of Statutory Health Insurances (GKV-SV). Before being imported into the research database, the data elements are checked for data format, completeness and plausibility. After each update of the research database, various counts are compared with the previous update to verify completeness. Due to the anonymity of the database, direct validation of the data (e.g., using medical records as the gold standard) is not possible.
10	Linkage	Due to the anonymization of the source data, linkage is not possible.
11	Vital status	The date of death is recorded as the last day of the quarter in which the death occurred (i.e., 30/31st of Mar/Jun/Sept/Dec) as reported to the health insurance (no linkage to death registry). The cause of death is not available.
12	Limitations	Ambulatory diagnoses and procedures are summarised in the source on a quarterly basis. Both are mapped to the observation table with the date set to the last day of the respective quarter (i.e., 30/31st of Mar/Jun/Sept/Dec) and the concept “History of event within 3 months” (observation_concept_id 1340222), with the actual diagnosis or procedure concept_id recorded in the field “value_as_concept_id”. There is no vocabulary for the German pharmaceutical product codes (PZN). A direct source-to-standard-mapping has been done manually by InGef but is incomplete. The drug exposure duration is unknown. Following OMOP conventions, the end date is always set to dispensing date + 29. Outpatient and inpatient procedures are recorded as OPS codes (German Procedure Classification), for which the vocabulary is incomplete. Approx. 10.5 Million insurees are

#	Section	Description
		included in the database, 7.8 Million of these actively insured in 2024. This corresponds to 7% of the total German population. Data are longitudinally linked over a period of currently ten years.
13	Main references	Andersohn F, Walker J "Characteristics and external validity of the German Health Risk Institute (HRI) Database." Pharmacoepidemiology and drug safety (2016): 26530279
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111207 Website: https://www.ingef.de/en/

Germany, IQVIA Disease Analyser (IQVIA DA Germany)

#	Section	Description
1	Data source identification and country	IQVIA DA Germany (IQVIA Disease Analyzer Germany) Germany
2	Data partner information section	IQVIA
3	Coverage and timespan	Data collection since: 1989 Extent: Nation-wide. GP and specialists in Germany using specific patient management software.
4	Healthcare setting / type of data	Primary care – General Practioner, and primary care specialists (e.g. paediatricians). Diagnoses, medication, and procedures from an ambulatory setting. Medications are recorded as prescriptions of marketed products.
5	Data collection process	Outpatient electronic health records. By clinicians at healthcare contact.
6	General representativeness	No specific details on general representativeness given.
7	Data content /source coding	Prescription is on product code level (German PZN), ICD10, NFC, Local lab coding.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. There can be patients registered under different ID numbers, because there is no linkage between different GPs.
9	Quality control (data source specific)	Data is quality checked on plausibility.
10	Linkage	No.
11	Vital status	Death information is derived from medical events.
12	Limitations	No database-specific limitations documented. General limitations for the data type applicable.
13	Main references	No main reference provided.
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/104282 Website: https://www.iqvia.com/

The Netherlands, The Integrated Primary Care Information (IPCI)

#	Section	Description
1	Data source identification and country	IPCI (Integrated Primary Care Information) The Netherlands
2	Data partner information section	Erasmus University Medical Center Department of Medical Informatics
3	Coverage and timespan	Data collection since: 2006 Extent: Nation-wide. IPCI is a Dutch database that contains patient records from 2006 onwards. However, it mainly covers the central part of the country, including the most densely populated area (the 'Randstad') and non-urban areas. IPCI contains information on all patients registered with GPs responsible for non-emergency care and referrals. A patient is registered at birth or at first encounter with the GP.
4	Healthcare setting / type of data	Primary care – General Practitioner. Data is collected from primary care EHR. This includes demographic information, complaints and symptoms, diagnoses, laboratory test results, lifestyle factors (in limited amount), and correspondence with secondary care, such as referral and discharge letters.
5	Data collection process	Outpatient electronic health records. Data is entered into the EHR system by the GPs, during or after the visit. The patientdossiers are collected by Erasmus MC data managers and combined in one harmonized database. Several checks are done on this database to ensure correct data processing. Persons can have dossiers at multiple GPs.
6	General representativeness	More than 99% of the Dutch population has health insurance, and almost all citizens are registered with a general practitioner. Over 12 months, around 78% of the population has at least one contact with their GP. IPCI included around 350 GP practices out of around 5000 in the country (~ 7%). The demographic composition of the IPCI population mirrors that of the general Dutch population in terms of age and sex.
7	Data content /source coding	Dutch GPs use mainly Dutch standard codes, like ICPC-1 and Diagnostische Bepalingen maintained by NHG. And for therapy the G-Standard is used, maintained by ZIndex.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. Patients can be registered under different IDs, but since a patient can only be registered at one GP at a time, the observations periods will not overlap.
9	Quality control (data source specific)	Prior to each data release, extensive quality control steps are performed, e.g., comparison of patient characteristics between practices, and checks to identify abnormal temporal data patterns in practices. For each practice, around 200 quality indicators are obtained. Of these indicators, a quarter refer to population characteristics, e.g. number of birth and mortalities relative to practice size, temporal consistency. The other indicators are based on medical data, e.g. distribution of measurement values, frequencies of diagnoses and procedures relative to age, completeness of data. The indicators are combined in a couple of quality scores for each practice. For these scores, cut-off values for acceptable quality have been defined. Practices with a score below a cut-off are excluded for research. This approach has shown to be very important, for example to check if data from practices that just joined the database are at an acceptable level of quality. The details of the approach, like the cut-off values for acceptance, are based on years of experience. In addition, trends are compared with the previous database release. Extensive quality control steps are performed before each data release. These include comparing patient characteristics between practices and checks to identify abnormal temporal data patterns in practices. Additional checks include over 200 indicators related to population characteristics (e.g., reliability of birth and mortality rates) and medical data (e.g., availability of durations of prescriptions and completeness of laboratory results). Records of low quality are excluded from the database.

#	Section	Description
10	Linkage	Linkage requires additional approval steps and needs to be assessed on a case-by-case basis. IPCI is not routinely linked with other databases.
11	Vital status	Vital status (death date and cause) is collected based on GP records.
12	Limitations	The main limitation comes with the fact that IPCI is limited to GP records, and although it contains information on referrals and discharge letters, it may not fully capture specific hospital information. IPCI does not include coded/detailed data about medications/procedures/test results from the hospital or other care-providers.
13	Main references	de Ridder MAJ, de Wilde M,de Ben C,Leyba AR,Mosseveld BMT,Verhamme KMC,van der Lei J,Rijnbeek PR "Data Resource Profile: The Integrated Primary Care Information (IPCI) database, The Netherlands." International journal of epidemiology (2022): 35182143
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/42618 Website: http://www.ipci.nl

Clinical Practice Research Datalink GOLD, the United Kingdom (CPRD GOLD)

#	Section	Description
1	Data source identification and country	CPRD GOLD (Clinical Practice Research Datalink GOLD) The United Kingdom
2	Data partner information section	University of Oxford NDORMS
3	Coverage and timespan	Data collection since: 1987 Extent: Nation-wide. CPRD GOLD consists of patients in contributing practices using Vision software. Historically this covered the whole of the UK, but the number of contributing practices in the England is dropping. In January 2025 only 3 practices from England were a part of CPRD GOLD, while historical patient data were from the whole of the UK, and will continue to be so. In the future, no practices from England will be present, only practices from Scotland, Wales, and Northern Ireland.
4	Healthcare setting / type of data	Primary care – General Practitioner, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. CPRD GOLD data include patient demographics, biological measurements, clinical symptoms and diagnoses, referrals to specialist/hospital and their outcome, laboratory tests/results, and prescribed medications.
5	Data collection process	Outpatient electronic health records. Data are entered by clinicians into the EHR. Data is processed by CPRD that provides data releases for research.
6	General representativeness	In the last 10 years, the CPRD GOLD regional distribution of currently contributing general practitioner (GP) practices has significantly shifted, resulting in many new practices joining from Scotland, Wales, and Northern Ireland, and fewer participating from England. These changes have affected the CPRD GOLD population size, regional coverage, and eligibility for data linkages. CPRD GOLD January 2024 contains >21.3 million historical and current patients (12.9 in England, 3.1 in Wales, 4.7 in Scotland, 0.7 in Northern Ireland). Of these, nearly 3 million are currently registered in a GP practice and represent ~4.3% of the estimated current UK population (0.1% in England, 32.3% in Wales, 28.6% in Scotland, 16.2% in Northern Ireland). Patients currently registered in CPRD GOLD January 2024 are broadly representative of the UK population with respect to age and sex. Reference: https://doi.org/10.1093/ije/dyaf077

#	Section	Description
7	Data content /source coding	Gemscript, Read, dm+d
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. In GOLD, a patient can be registered under different ID numbers upon changing practice or re-registration. Researchers are not able to identify these patients, as the data are anonymised. However, GOLD covers less than 5% of the current UK GP practices and it is unlikely that an individual who does change GP practice ends up in another GP practice which uses the Vision software and accepts the CPRD data collection agreement. The very small number of duplicated IDs will have different observation periods and should not have an impact on the data analyses.
9	Quality control (data source specific)	CPRD GOLD only includes practices whose data quality is assessed to be up-to-standard (UTS). Each practice is associated to an UTS date set when the data quality standards become satisfactory, and CPRD recommend using only longitudinal data starting from this UTS date. Every time CPRD collect the EHR from a practice, checks are run for the data quality standards, and if they are not adequate, the EHR is not accepted. When the data quality becomes acceptable again, CPRD updates the practice UTS date. CPRD also checks data quality standards at the patient level, and associates each patient with a flag, reporting if its data are acceptable for clinical research. Only patients with acceptable data quality are included in the population to be mapped to CDM.
10	Linkage	CPRD GOLD can be linked to several sources, however our Oxford OMOP CDM is only linked to the CPRD GOLD Ethnicity Record and to the CPRD Townsend Deprivation Index at the Practice Level.
11	Vital status	The date of death in CPRD GOLD has been validated against the Population registry (ONS) mortality data. Reference: https://doi.org/10.1002/pds.4747
12	Limitations	The main limitation is due to the fact that CPRD GOLD is limited to GP records, and although it contains information on referrals and discharge letters, it may not fully capture specific hospital information. Events from hospital and specialist care are not covered.
13	Main references	Sanchez-Santos MT, Axson EL,Dedman D,Delmestri A "Data Resource Profile Update: CPRD GOLD." International journal of epidemiology (2025): 40499193
14	Link to HMA-EMA catalogue and data source webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111113 Website: https://www.cprd.com/data/primary-care-data/cprd-gold

ANNEX II. Fitness for use assessment

Croatia, National Health Information System (NAJS)

NAJS was included in this study because it is a primary care, outpatient specialist care, and inpatient care registries data source that provides relevant information on AD in the general adult population (≥18). The CDM population comprises all publicly insured persons residing in Croatia. The inclusion of NAJS enhances the geographical diversity of data sources, with adequate data coverage over the study period.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in NAJS was 28,900 individuals.

Moreover, data availability and follow-up in NAJS is sufficient, as accurate data availability starts in 2017 and the date of most recent data extraction is 02/08/2025, which aligns with the study period. The median follow-up of the first observation period in NAJS is 3.84k (IQR: 3.31k–3.93k).

There were no study specific limitations present in NAJS. Due to data availability, the study period for NAJS was from 01/01/2017 to 31/12/2024. Lastly, IRB approval for NAJS was estimated to take one month, which made the execution of this study feasible within the current study timelines.

Denmark, Danish Data Health Registries (DK-DHR)

DK-DHR was included in this study because it is a nationwide registry data source, comprising inpatient, outpatient, and emergency care data from hospitals that provides relevant information on AD, radiology, medications, and relevant comorbidities in the general adult population (≥18). The inclusion of DK-DHR enhances the geographical diversity of data sources, with adequate data coverage over the study period.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in DK-DHR was 107,900 individuals.

Moreover, data availability and follow-up in DK-DHR is sufficient, as data availability starts from 01/01/1995 and the date of most recent data extraction is 10/04/2025, which aligns with the study period. The median follow-up of the first observation period in DK-DHR is 7.92k (IQR: 2.61k–10.90k).

There were no study specific limitations present in DK-DHR. Due to data availability, the study period for DK-DHR was from 01/01/2014 to 31/12/2024. Lastly, DK-DHR operates under blanket approval, which made the execution of this study feasible within the current study timelines.

Germany, InGef Research Database (InGef RDB)

InGef RDB was included in this study because it includes primary care, hospital inpatient care, and secondary outpatient care claims data, which provides relevant information on AD, radiology, medications, and relevant comorbidities in the general adult population (≥18). The CDM population comprises individuals from all regions in Germany who are covered by one of the approximately 50 contributing statutory health insurance companies, ~15% of the German population. The inclusion of InGef RDB enhanced the geographical diversity of data sources.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in InGef RDB was 27,500 individuals.

Moreover, data availability and follow-up in InGef RDB is sufficient, as data availability starts from 01/01/2015 and the date of most recent data extraction is 18/04/2025, which aligns with the study period. The median follow-up of the first observation period in InGef RDB is 3.56k (IQR: 1.40k–3.65k).

There were no specific limitations present in InGef RDB. Due to data availability, the study period for InGef RDB was from 01/01/2015 to 31/12/2024. Lastly, InGef RDB operates under blanket approval, which made the execution of this study feasible within the current study timelines.

Germany, IQVIA Disease Analyser (IQVIA DA Germany)

IQVIA DA Germany was included in this study because it includes primary and outpatient secondary care electronic health record data that provides relevant information on AD, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18). The inclusion of IQVIA DA Germany enhanced the geographical diversity of data sources.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in IQVIA DA Germany was 79,300 individuals.

Moreover, data availability and follow-up in IQVIA QA Germany is sufficient, as data availability starts from 01/01/1992 and the date of most recent data extraction is 10/04/2025, which aligns with the study period. The median follow-up of the first observation period in IQVIA DA Germany is 116 (IQR: 0–1.61k).

There were no specific limitations present in IQVIA DA Germany. Due to data availability, the study period for IQVIA DA Germany was from 01/01/2014 to 31/12/2024. Lastly, IQVIA QA Germany operates under blanket approval, which made the execution of this study feasible within the current study timelines.

The Netherlands, The Integrated Primary Care Information (IPCI)

IPCI was included in this study because it includes primary care electronic health record data that provides relevant information on AD, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18). The inclusion of IPCI enhanced the geographical diversity of data sources.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in IPCI was 13,400 individuals.

Moreover, data availability and follow-up in IPCI is sufficient, as data availability starts from 01/01/2006 and the date of most recent data extraction is 16/04/2025, which aligns with the study period. The median follow-up of the first observation period in IPCI is 1.73k (IQR: 791–3.07k).

There were no specific limitations present in IPCI. Due to data availability, the study period for IPCI will be from 01/01/2014 to 31/12/2024. Lastly, IRB approval for IPCI was obtained within approximately one month, which made the execution of this study feasible within the current study timelines.

The United Kingdom, Clinical Practice Research Datalink GOLD (CPRD GOLD)

CPRD GOLD was included in this study because it includes primary care-GP, primary specialist care, secondary specialist care of outpatient electronic health records, which provides relevant information on AD, radiology, activities of daily living, caregiver support, medications, and relevant comorbidities in the general adult population (≥18). The CDM population comprises all persons residing in the United Kingdom. The inclusion of CPRD GOLD enhanced the geographical diversity of data sources.

Based on a preliminary feasibility assessment, the expected number of person-counts for AD in CPRD GOLD was 87,500 individuals.

Moreover, data availability and follow-up in CPRD GOLD is sufficient, as data availability starts from 01/01/1988 and the date of most recent data extraction is 15/03/2025, which aligns with the study period. The median follow-up of the first observation period in CPRD GOLD is 2.15k (IQR: 728–4.94k).

There were no specific limitations present in CPRD GOLD. Due to data availability, the study period for CPRD GOLD will be from 01/01/2014 to 31/12/2024. Lastly, CPRD GOLD operates under blanket approval, which made the execution of this study feasible within the current study timelines.

EMA Data Quality Framework for EU medicines regulation: application to Real-World Data for more information (https://www.ema.europa.eu/system/files/documents/other/data-quality-framework-eu-medicines-regulation-application-real-world-data_en.pdf).

ANNEX III. Operational and reporting considerations

Table S1. List of conditions definitions.

Phenotype	Concept name	Concept id (including descendants)	Exclude concept id	Vocabulary
Alzheimer's disease	Alzheimer's disease	378419	-	SNOMED
	Descendant concepts of Alzheimer's disease: *Altered behavior in Alzheimer's disease, Alzheimer disease with psychosis, Alzheimer's disease co-occurrent with delirium, Alzheimers continuum, Autosomal dominant Alzheimer disease due to mutation of amyloid precursor protein, Autosomal dominant Alzheimer disease due to mutation of presenilin 1, Autosomal dominant Alzheimer disease due to mutation of presenilin 2, Behavioral disturbance co-occurrent and due to late onset Alzheimer dementia, Delusions in Alzheimer's disease, Dementia of the Alzheimer type with behavioral disturbance, Depressed mood in Alzheimer's disease, Early Alzheimers disease, Early onset Alzheimer's disease with behavioral disturbance, Familial Alzheimer's disease of early onset, Familial Alzheimer's disease of late onset, Focal Alzheimer's disease, Frontal variant non-amnestic Alzheimer disease, High level Alzheimers neuropathology changes, Intermediate level Alzheimers neuropathology changes, Logopenic non-amnestic Alzheimer disease, Low level Alzheimers neuropathology changes, Mixed dementia, Non-amnestic Alzheimer disease, Non-familial Alzheimer's disease of early onset, Non-familial Alzheimer's disease of late onset, Primary degenerative dementia of the Alzheimer type, presenile onset, Primary degenerative dementia of the Alzheimer type, presenile onset in remission, Primary degenerative dementia of the Alzheimer type, presenile onset, uncomplicated, Primary degenerative dementia of the Alzheimer type, presenile onset, with delirium, Primary degenerative dementia of the Alzheimer type, presenile onset, with delusions, Primary degenerative dementia of the Alzheimer type, presenile onset, with depression, Primary degenerative dementia of the Alzheimer type, senile onset, Primary degenerative dementia of the Alzheimer type, senile onset in remission, Primary degenerative dementia of the Alzheimer type, senile onset, uncomplicated, Primary degenerative dementia of the Alzheimer type, senile onset, with behavioral disturbance, Primary degenerative dementia of the Alzheimer type, senile onset, with delirium, Primary degenerative dementia of the Alzheimer type, senile onset, with delusions, Primary degenerative dementia of*	*Descendant concept IDs:* *44784643, 37166126, 37395572, 3170377, 608051, 603149, 608060, 37117145, 44782726, 43530664, 44782727, 3179057, 44782432, 4043241, 4043243, 4043377, 37167060, 3184947, 3176418, 37167043, 3177168, 43021816, 36716558, 4043242, 4043244, 4218017, 44782941, 4277444, 4277746, 4182539, 4019705, 4220313, 44782940, 4278830, 762578, 4167839, 4204688, 4097384, 1340510, 4043379.*	-	*SNOMED*

Phenotype	Concept name	Concept id (including descendants)	Exclude concept id	Vocabulary
	the Alzheimer type, senile onset, with depression, Progression of Alzheimer's disease, Progressive aphasia in Alzheimer's disease.			
Down’s syndrome	Complete trisomy 21 syndrome, Partial trisomy 21 in Down's syndrome,	439125, 4110269	-	SNOMED
Stroke	Normal cardiac stroke volume, Decreased cardiac stroke volume, Stroke test finding, Ischemic stroke, Increased cardiac stroke volume, Haemorrhagic stroke, Autosomal recessive leukoencephalopathy, ischemic stroke, retinitis pigmentosa syndrome, Acute stroke, Late effects of cerebral ischemic stroke, Cerebellar stroke, Stroke co-occurrent with migraine, Epilepsy due to perinatal stroke, Decreased cardiac stroke volume index, Cerebral ischemic stroke due to subarachnoid hemorrhage, Cerebral ischemic stroke due to hypercoagulable state, Cerebral ischemic stroke due to global hypoperfusion with watershed infarct, Cerebral ischemic stroke due to dissection of artery, Seizure disorder as sequela of stroke, Sequela of lacunar stroke, Sequela of cardioembolic stroke, Weakness of face muscles as sequela of stroke, Sequela of thrombotic stroke, Weakness of extremities as sequela of stroke, Occlusion of cerebral artery with stroke, Spasticity as sequela of stroke, Alteration of sensation as late effect of stroke, Weakness as a late effect of stroke, Hemiplegia and/or hemiparesis following stroke	372654, 440426, 761790, 4006295, 4023571, 4045755, 4046363, 4079008, 4088120, 4099974, 4111710, 4153352, 4159140, 4159152, 4168056, 4181404, 4219010, 4236498, 4243337, 4280420, 4310996, 4338810, 35609033, 36675148, 36684840, 36716860, 36716999, 37110241, 37110521, 37118679, 37312013, 37312014, 37312015, 37312017, 43530665, 43530732, 43530736, 43530744, 43531592, 43531595, 43531605, 43531610, 43531617, 44782753, 44782781	3663227, 37110521, 36684840	SNOMED
Atrial fibrillation	Atrial fibrillation	313217	-	SNOMED
Myocardial infarction	Myocardial infarction	4329847	-	SNOMED
Heart failure	Heart failure	316139	-	SNOMED
Hypertension	Hypertensive disorder, Hypertension secondary to endocrine disorder, Hypertension secondary to kidney transplant, Rebound hypertension, Secondary diastolic hypertension, Secondary hypertension, Transient hypertension, Transient hypertension of pregnancy, Transient hypertension of pregnancy - not delivered, Hypertensive complication, Hypertensive renal disease, Hypertensive heart failure, Hypertensive heart AND renal disease	316866, 4110948, 4178312, 4221991, 4253928, 319826, 4199306, 441922, 136760, 42709887, 201313, 444101, 195556	4071202, 762994, 43021830, 137940, 141639, 4062906, 42599748	SNOMED
Diabetes	Diabetes mellitus, Complication due to diabetes mellitus	201820, 442793	-	SNOMED
Hypercholesterolemia	Pure hypercholesterolemia, Hypercholesterolemia	437827, 4029305	-	SNOMED
Hypertriglyceridemia	Hypertriglyceridemia	4120314	-	SNOMED

Phenotype	Concept name	Concept id (including descendants)	Exclude concept id	Vocabulary
Brain MRI	Brain MRI measuring method, MRI for measurement of brain volume, MRI of brain, MRI of brain and brain stem without contrast, MRI of brain and cervical spinal cord without contrast, MRI of brain and facial bones without contrast, MRI of brain and internal auditory canal without contrast, MRI of brain without contrast, MRI of cerebrospinal fluid flow without contrast, MRI of head, Magnetic Resonance Imaging (MRI) of Brain using Other Contrast, Magnetic Resonance Imaging (MRI) of Brain using Other Contrast, Unenhanced and Enhanced, Magnetic Resonance Imaging (MRI) of Sella Turcica/Pituitary Gland using Other Contrast, Magnetic Resonance Imaging (MRI) of Sella Turcica/Pituitary Gland using Other Contrast, Unenhanced and Enhanced, Magnetic resonance imaging for measurement of brain volume with contrast, Magnetic resonance spectroscopy of brain without contrast	35810878, 44784284, 37311324, 36713175, 36713041, 36713233, 36713228, 36713262, 36713055, 4082979, 2789360, 2789359, 2789363, 2789362, 44784285, 36713050		SNOMED
Brain PET-F18	PET CT of brain, PET CT of brain using 18F-FDOPA (fluorodopa 18-F), PET+CT Brain for amyloidosis, PET+CT Brain for tau protein, Positron Emission Tomographic (PET) Imaging of Brain using Fluorine 18 (F-18), Positron emission tomography of brain using fluorodeoxyglucose (18-F), Positron emission tomography with computed tomography of brain using fluorodeoxyglucose (18-F), Positron emission tomography with computed tomography of brain using flutemetamol (18-F), Positron emission tomography of brain	37109091, 37310737, 36304731, 37021253, 2793369, 35622623, 35608074, 36713668, 35608075	-	SNOMED
Neuropsychiatric symptoms	Altered behavior in Alzheimer's disease, Alzheimer's disease co-occurrent with delirium, Behavioral disturbance co-occurrent and due to late onset Alzheimer dementia, Delusions in Alzheimer's disease, Dementia of the Alzheimer type with behavioral disturbance, Depressed mood in Alzheimer's disease, Early onset Alzheimer's disease with behavioral disturbance	44784643, 37395572, 37117145, 44782726, 43530664, 44782727, 44782432	-	SNOMED
Alterations in activities of daily living	Activity of daily living (ADL) alteration, Assisting with activity of daily living, Finding related to ability to perform personal hygiene activity, Functionally dependent, Instrumental activity of daily living, Physical functional dependency, Unfit for activity	4032520, 4128088, 4274102, 36713755, 4044726, 4030753, 44811145	4216322, 4114949, 4108026, 4110603, 4118377, 4108179, 4116022, 4115872, 4115873, 4129511, 4112904, 4116884, 4113964, 4108180, 4101952, 37018837, 4108027, 4113310, 4115878, 4114936, 4113877, 4120240, 4115982, 4115954, 4116034, 4115976, 4127369, 4104322,	SNOMED

Phenotype	Concept name	Concept id (including descendants)	Exclude concept id	Vocabulary
			4113314, 4116019, 4115074, 4108022, 4128345, 4128695, 4128696, 4116027, 4114728, 4108029, 4116038, 4113861, 4115892, 4110921, 4111047, 4110626, 4115080, 4116029, 4114814, 4109997, 4110606, 4108048, 4120246, 4113177, 4113179, 4110592, 4108024, 4108178, 4116021, 4101953, 4101956, 4129510, 4101951, 4111693, 4109857, 4108050, 4114932, 4114935, 4111983, 4111977, 4111988, 4116033, 4120747, 4115868, 4108175, 4114956, 4101972, 4113866, 4108021, 45767126, 4126101, 4125452, 4114962, 4113864, 4111974, 4110472, 4102101, 4113860, 4111981, 4115891, 4114745, 4113326, 4114964, 4104330, 4108046, 4108045, 4109998, 4116105, 4113175, 4109987, 4110599, 4120228, 4113168	
Caregiver support	Caregiver support, Problem related to life management difficulty, Patient dependence on care provider, Need for personal care assistance, Support, Minimum level of support, Maximum level of support	4303295, 43020485, 4022076, 42535090, 4055254, 44791823, 44791822	44790457, 44792004, 607715, 46272455, 37164586, 36715096, 46273902, 37164347, 37170022, 45769206, 37311826, 37163172, 37170021, 35609662, 36685469, 4254219	SNOMED
MCI before AD diagnosis	Mild cognitive impairment review, Minimal cognitive impairment, Mild neurocognitive disorder	37398899, 439795, 4297400	-	SNOMED

MRI=Magnetic Resonance Imaging, PET-F18=Positron Emission Tomography with Fluorine-18, CT=Computed tomography, ADL=Activity of daily living, MCI=Mild Cognitive Impairment, AD=Alzheimer's disease

Table S2. List of medicines definitions.

Substance Name	Concept name	Class	Ingredient Concept ID	Include descendants
Alzheimer disease drugs				
Memantine	Memantine	Ingredient	701322	Yes
Donepezil	Donepezil	Ingredient	715997	Yes
Rivastigmine	Rivastigmine	Ingredient	733523	Yes
Galantamine	Galantamine	Ingredient	757627	Yes
Hypertension drugs				
Antihypertensives:				
ACE inhibitors (lisinopril, enalapril)	lisinopril, enalapril	Ingredient	1308216, 1341927	Yes
ARBs (losartan, valsartan)	losartan, valsartan	Ingredient	1367500, 1308842	Yes
Beta-blockers (metoprolol, atenolol)	metoprolol, atenolol	Ingredient	1307046, 1314002	Yes
Calcium channel blockers (amlodipine, diltiazem)	amlodipine, diltiazem	Ingredient	1332418, 1328165	Yes
Diuretics (hydrochlorothiazide, furosemide, spironolactone)	hydrochlorothiazide, furosemide, spironolactone	Ingredient	974166, 956874, 970250	Yes
Stroke / Myocardial infarction / Atrial fibrillation / Heart failure drugs				
Antiplatelets:				
aspirin	aspirin	Ingredient	1112807	Yes
clopidogrel	clopidogrel	Ingredient	1322184	Yes
ticagrelor	ticagrelor	Ingredient	1116632	Yes
prasugrel	prasugrel	Ingredient	40163718	Yes
Anticoagulants/antithrombotics:		Ingredient		Yes
warfarin	warfarin	Ingredient	1310149	Yes
heparin	heparin	Ingredient	1367571	Yes
DOACs (apixaban, rivaroxaban, dabigatran, edoxaban)	apixaban, rivaroxaban, dabigatran etexilate, edoxaban	Ingredient	43013024, 40241331, 40228152, 45892847	Yes
fondaparinux	fondaparinux	Ingredient	1315865	Yes
bivalirudin	bivalirudin	Ingredient	19084670	Yes
Rhythm control drugs:				
amiodarone	amiodarone	Ingredient	1309944	Yes
flecainide	flecainide	Ingredient	1354860	Yes
propafenone	propafenone	Ingredient	1353256	Yes
sotalol	sotalol	Ingredient	1370109	Yes
dronedarone	dronedarone	Ingredient	40163615	Yes
dofetilide	dofetilide	Ingredient	1362979	Yes

Substance Name	Concept name	Class	Ingredient Concept ID	Include descendants
Heart failure treatments:				
ACE inhibitors (lisinopril, enalapril, ramipril, captopril, fosinopril, perindopril, quinapril, trandolapril, benazepril, moexipril, cilazapril)	lisinopril, enalapril, ramipril, captopril, fosinopril, perindopril, quinapril, trandolapril, benazepril, moexipril, cilazapril	Ingredient	1308216, 1341927, 1334456, 1340128, 1363749, 1373225, 1331235, 1342439, 1335471, 1310756, 19050216	Yes
ARBs (losartan, valsartan, candesartan, telmisartan)	losartan, valsartan, candesartan, telmisartan	Ingredient	1367500, 1308842, 1351557, 1317640	Yes
beta-blockers (metoprolol, bisoprolol, carvedilol, nebivolol)	metoprolol, bisoprolol, carvedilol, nebivolol	Ingredient	1307046, 1338005, 1346823, 1314577	Yes
diuretics (hydrochlorothiazide, furosemide, spironolactone, bumetanide, torsemide, metolazone, chlorothiazide, amiloride, chlorthalidone, indapamide, triamterene, eplerenone)	hydrochlorothiazide, furosemide, spironolactone, bumetanide, torsemide, metolazone, chlorothiazide, amiloride, chlorthalidone, indapamide, triamterene, eplerenone	Ingredient	974166, 956874, 970250, 932745, 942350, 907013, 992590, 991382, 1395058, 978555, 904542, 1309799	Yes
mineralocorticoid receptor antagonists (spironolactone, eplerenone)	spironolactone, eplerenone	Ingredient	970250, 1309799	Yes
SGLT2 inhibitors (dapagliflozin, empagliflozin)	dapagliflozin, empagliflozin	Ingredient	44785829, 45774751	Yes
sacubitril/valsartan	sacubitril / valsartan Oral Tablet	Ingredient	46275724	Yes, except 46275719, 1308842
ivabradine	ivabradine	Ingredient	46234437	Yes
vericiguat	vericiguat	Ingredient	739665	Yes
digoxin	digoxin	Ingredient	1326303	Yes
hydralazine/ isosorbide dinitrate	hydralazine/ isosorbide dinitrate Oral Tablet	Ingredient	40140795	Yes
Diabetes drugs				
Insulin	Insulins and analogues	Ingredient	21600713	Yes
Oral glucose-lowering drugs:		Ingredient		
Metformin (biguanide)	biguanide	Ingredient	1593986	Yes
Sulfonylureas (glipizide, gliclazide)	glipizide, gliclazide	Ingredient	1560171, 19059796	Yes
DPP-4 inhibitors (sitagliptin, linagliptin)	sitagliptin, linagliptin	Ingredient	1580747, 40239216	Yes
SGLT2 inhibitors (dapagliflozin, empagliflozin)	dapagliflozin, empagliflozin	Ingredient	44785829, 45774751	Yes
GLP-1 receptor agonists (liraglutide, semaglutide)	liraglutide, semaglutide	Ingredient	40170911, 793143	Yes
Hypercholesterolemia / Hypertriglyceridemia drugs				

Substance Name	Concept name	Class	Ingredient Concept ID	Include descendants
Lipid-lowering drugs:				
Statins (atorvastatin, simvastatin, rosuvastatin)	atorvastatin, simvastatin, rosuvastatin	Ingredient	1545958, 1539403, 1510813	Yes
Fibrates (fenofibrate, gemfibrozil)	fenofibrate, gemfibrozil	Ingredient	1551803, 1558242	Yes
Ezetimibe	Ezetimibe	Ingredient	1526475	Yes
PCSK9 inhibitors (alirocumab, evolocumab)	alirocumab, evolocumab	Ingredient	46275447, 46287466	Yes
Omega-3 fatty acids (for triglyceride lowering)	Omega-3 fatty acids	Ingredient	19106973	Yes
Down's syndrome drugs				
No disease-specific chronic medications	-	-	-	-

ACE=Angiotensin-converting enzyme, ARB=Angiotensin II receptor blockers, DOAC=Direct oral anticoagulants, DPP-4=dipeptidyl peptidase 4, SGLT2=Sodium-glucose cotransporter-2, GLP-1=Glucagon-like peptide-1, PCSK9=Proprotein convertase subtilisin/kexin type 9

ANNEX IV: Operational and reporting considerations

DATA MANAGEMENT

Data management

All data sources had previously mapped their data to the OMOP common data model. This enabled the use of standardised analytics and using DARWIN EU® tools across the network, since the structure of the data and the terminology system is harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org.

The analytic code for this study was written in R and used standardized analytics wherever possible. Each data partner executed the study code against their data source containing patient-level data and then returned the results (csv files), which only contained aggregated data. The results from each of the contributing data sites were then combined in tables and figures for the study report.

Data storage and protection

For this study, participants from various EU member states processed personal data from individuals that was collected in national/regional electronic health record data sources. Due to the sensitive nature of this personal medical data, it was important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All data sources used in this study were already used for pharmaco-epidemiological research and had a well-developed mechanism to ensure that European and local regulations dealing with ethical use of the data and adequate privacy control were adhered to. In agreement with these regulations, rather than combining person level data and performing only a central analysis, local analyses were, which generated nonidentifiable aggregate summary results.

The output files were stored in the DARWIN EU® Digital Research Environment (DRE). These output files did not contain any data that allowed identification of subjects included in the study. The DRE implemented further security measures to ensure a high level of stored data protection to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/20161 in the various member states.

QUALITY CONTROL

Data source quality control

When defining drug cohorts, non-systemic products were excluded from the list of included codes summarised on the ingredient level.

When defining cohorts for indications, a systematic search of possible codes for inclusion was identified using ATLAS. In addition, the CohortDiagnostics (https://github.com/OHDSI/CohortDiagnostics) R package was run to assess the use of different codes across the data sources contributing to the study and identify any codes potentially omitted in error.

The study code was based on DARWIN EU® R packages*: IncidencePrevalence* to estimate Incidence and Prevalence, and *CohortCharacteristics* and *PatientProfiles* for the population level characterisation, individual-level characterisation, and large-scale characterisation. These packages included numerous P4-C1-021 Study Protocol Version: V2.0 Dissemination level: Public 47/61 automated unit tests to ensure the validity of the codes, alongside software peer review and user testing. The R package is made publicly available via GitHub.

ANNEX V: Flowcharts of exclusion criteria for AD cohort

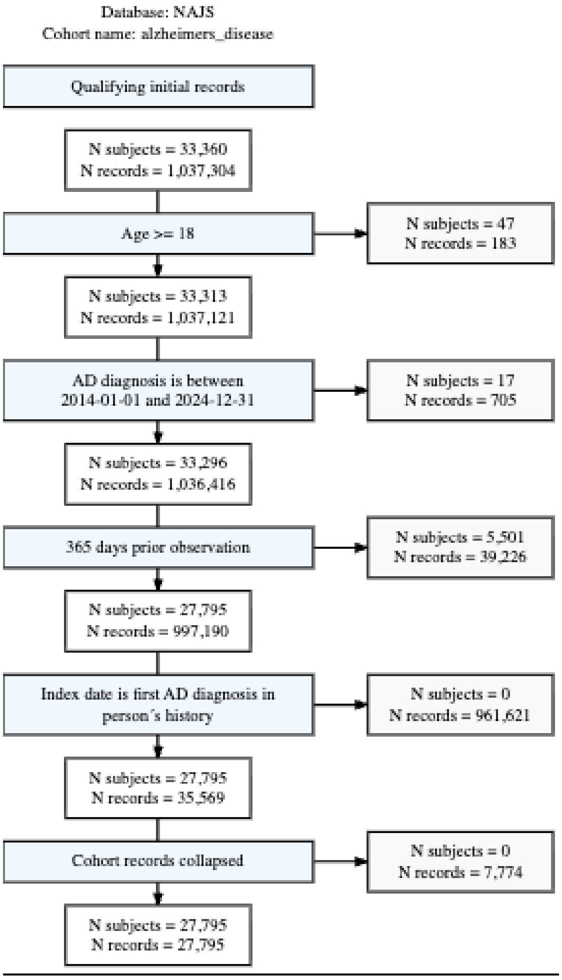

Figure S1. Flowchart of number of individuals with AD excluded in each exclusion criteria for NAJS.

NAJS=National Public Health Information System

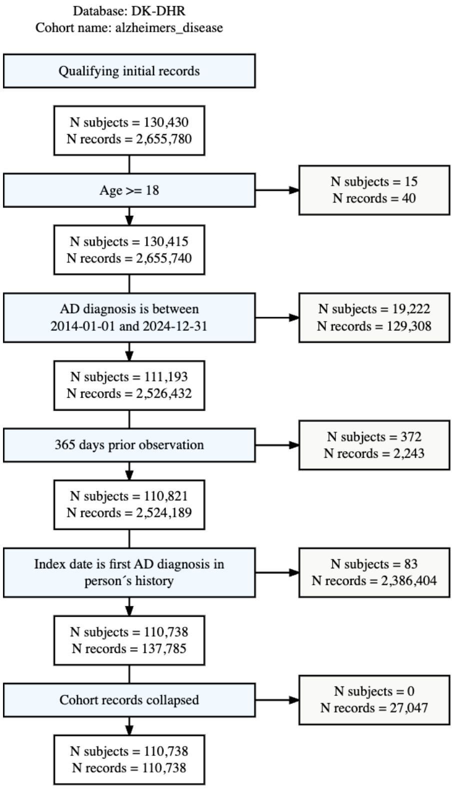

Figure S2. Flowchart of number of individuals with AD excluded in each exclusion criteria for DK-DHR.

DK-DHR=Danish Data Health Registries

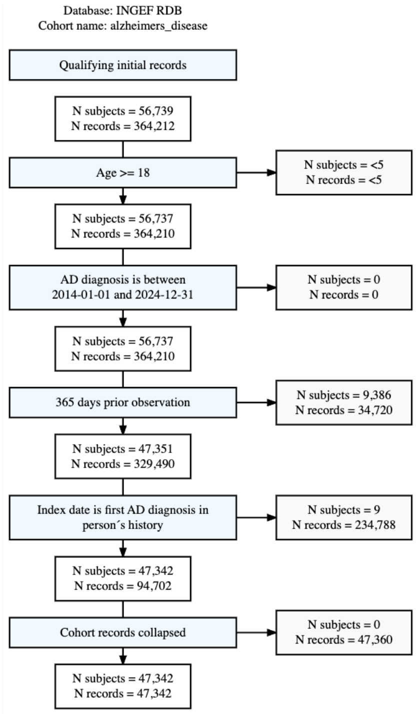

Figure S3a. Flowchart of number of individuals with AD excluded in each exclusion criteria for InGef RDB.

This figure includes inpatient diagnosis only.
InGef RDB=InGef Research Database

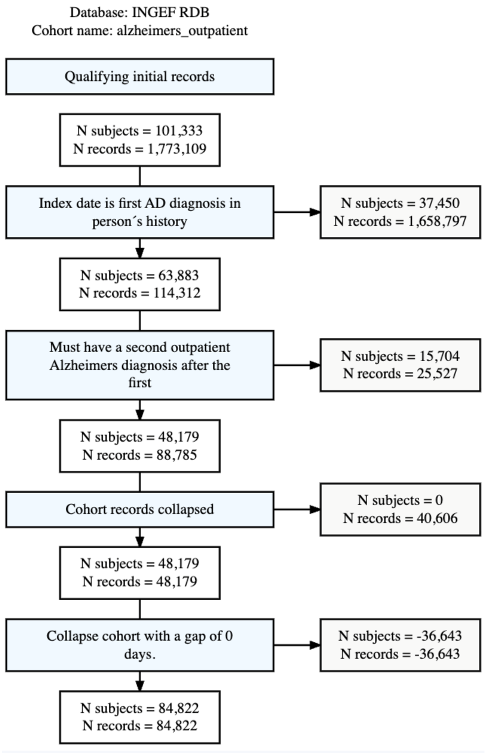

Figure S3b. Flowchart of number of individuals with AD excluded in each exclusion criteria for InGef RDB.

This figure includes the outpatient diagnosis, in which inpatient diagnosis are added at the end, giving a total of 84,822 individuals with AD.
InGef RDB=InGef Research Database

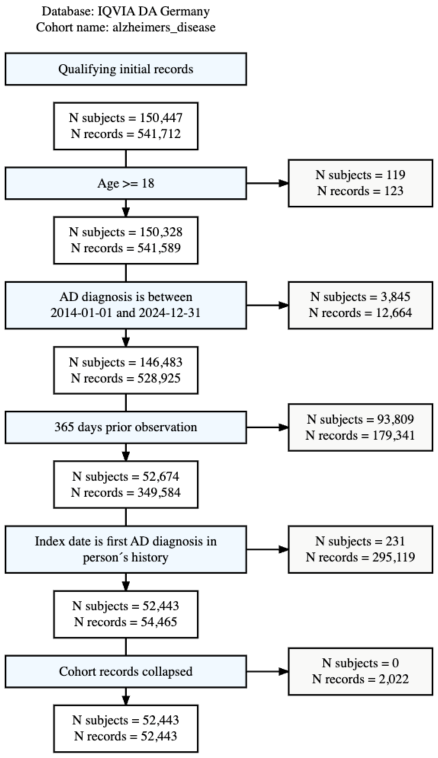

Figure S4. Flowchart of number of individuals with AD excluded in each exclusion criteria for IQVIA DA Germany.

IQVIA DA Germany=IQVIA Disease Analyzer Germany

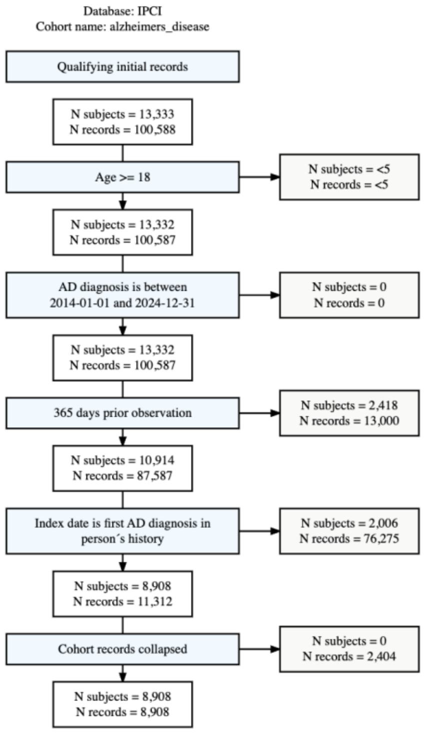

Figure S5. Flowchart of number of individuals with AD excluded in each exclusion criteria for IPCI.

IPCI=Integrated Primary Care Information

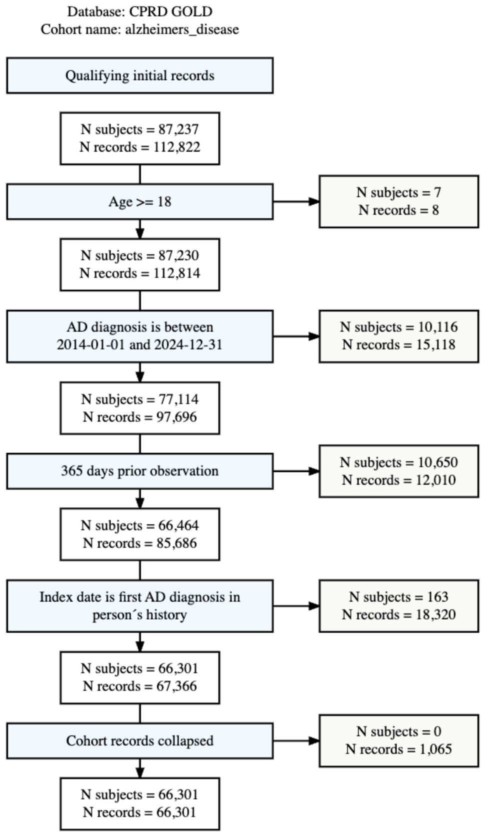

Figure S6. Flowchart of number of individuals with AD excluded in each exclusion criteria for CPRD GOLD.

CPRD=Clinical Practice Research Datalink

ANNEX VI: Incidence per 100,000 PY of AD in the general adult population

Table S1. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in NAJS stratified by calendar year, sex, and age categories.

	NAJS			
	Number of participants	Follow-up (PYs)	Number of individuals with AD	Incidence Rates/100,000 PYs (95% CI)
Calendar year				
2014	NA	NA	NA	NA
2015	NA	NA	NA	NA
2016	NA	NA	NA	NA
2017	NA	NA	NA	NA
2018	3,476,133	3,397,787.28	3,037	89.38 (86.23–92.62)
2019	3,508,007	3,435,173.35	3,075	89.52 (86.38–92.74)
2020	3,530,047	3,472,649.57	2,673	76.97 (74.08–79.95)
2021	3,570,033	3,487,034.94	2,870	82.31 (79.32–85.37)
2022	3,576,333	3,507,230.85	3,064	87.36 (84.30–90.51)
2023	3,588,238	3,524,276.88	2,880	81.72 (78.76–84.76)
2024	3,610,059	3,562,858.06	2,909	81.65 (78.71–84.67)
Sex				
Female	2,124,162	18,259,223.81	18,440	100.99 (99.54–102.46)
Male	2,019,080	16,286,589.51	9,351	57.42 (56.26–58.59)
Age categories				
Early-onset AD (18–65)	3,324,106	26,076,060.97	2,256	8.65 (8.30–9.02)
Late-onset AD (≥66)	1,390,143	8,469,752.35	25,535	301.48 (297.80–305.20)
18–55	2,690,557	20,109,397.27	632	3.14 (2.90–3.40)
56–65	1,223,413	5,966,663.70	1,624	27.22 (25.91–28.58)
66–75	1,012,459	4,678,088.36	6,850	146.43 (142.98–149.94)
76–85	669,663	2,883,636.83	13,430	465.73 (457.89–473.68)
≥86	259,588	908,027.16	5,255	578.73 (563.18–594.59)

NAJS=National Public Health Information System, AD=Alzheimer’s disease, PYs=Person-years, CI=Confidence Interval, NA=Not applicable

Table S2. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in DK-DHR stratified by calendar year, sex, and age categories.

	DK-DHR			
	Number of participants	**Follow-up (PYs)**	**Number of individuals with AD**	**Incidence Rates/100,000 PYs (95% CI)**
Calendar year				
2014	4,550,072	4,426,440.06	5,385	121.66 (118.43–124.95)
2015	4,597,995	4,470,306.99	5,477	122.52 (119.30–125.81)
2016	4,639,615	4,525,102.55	5,539	122.41 (119.20–125.67)
2017	4,674,892	4,548,121.70	5,713	125.61 (122.38–128.91)
2018	4,708,726	4,580,033.54	5,423	118.40 (115.28–121.60)
2019	4,734,907	4,608,903.20	5,542	120.25 (117.10–123.45)
2020	4,750,160	4,645,196.43	5,678	122.23 (119.08–125.46)
2021	4,780,549	4,661,806.95	5,845	125.38 (122.19–128.64)
2022	4,847,535	4,710,737.89	5,994	127.24 (124.04–130.50)
2023	4,887,520	4,754,483.26	6,883	144.77 (141.37–148.23)
2024	NA	NA	NA	NA
Sex				
Female	2,939,069	25,323,678.40	37,730	148.99 (147.49–150.50)
Male	2,909,393	24,703,993.41	25,761	104.28 (103.01–105.56)
Age categories				
Early-onset AD (18–65)	5,896,736	73,206,357.52	5,494	7.50 (7.31–7.71)
Late-onset AD (≥66)	2,119,164	19,557,160.00	105,243	538.13 (534.88–541.39)
18–55	4,190,857	30,817,541.63	629	2.04 (1.89–2.21)
56–65	1,569,530	7,795,896.51	2,284	29.30 (28.11–30.52)
66–75	1,306,471	6,702,573.36	13,725	204.77 (201.36–208.23)
76–85	830,013	3,663,669.74	31,807	868.17 (858.66–877.77)
≥86	289,534	1,047,990.58	15,046	1,435.70 (1,412.85–1,458.83)

DK-DHR=Danish Data Health Registries, AD=Alzheimer’s disease, PYs=Person-years, CI=Confidence Interval, NA=Not applicable

Table S3. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in InGef RDB stratified by calendar year, sex, and age categories.

	InGef RDB			
	Number of participants	**Follow-up (PYs)**	**Number of individuals with AD**	**Incidence Rates/100,000 PYs (95% CI)**
Calendar year				
2014	NA	NA	NA	NA
2015	NA	NA	NA	NA
2016	6,365,230	6,088,533.46	12,080	198.41 (194.88–201.98)
2017	6,375,546	6,132,687.10	10,955	178.63 (175.30–182.01)
2018	6,404,695	6,170,555.74	10,085	163.44 (160.26–166.66)
2019	6,442,368	6,193,395.35	9,822	158.59 (155.47–161.76)
2020	6,432,822	6,225,250.90	8,798	141.33 (138.39–144.31)
2021	6,483,382	6,218,759.37	9,120	146.65 (143.66–149.69)
2022	6,531,837	6,273,021.41	9,084	144.81 (141.85–147.82)
2023	6,595,560	6,335,591.08	9,064	143.06 (140.14–146.04)
2024	NA	NA	NA	NA
Sex				
Female	4,365,865	31,116,541.78	47,633	153.08 (151.71–154.46)
Male	4,470,761	30,646,134.54	37,188	121.35 (120.12–122.59)
Age categories				
Early-onset AD (18–65)	7,813,828	50,935,778.18	4,442	8.72 (8.47–8.98)
Late-onset AD (≥66)	1,801,071	10,826,898.15	80,379	742.40 (737.28–747.55)
18–55	6,817,835	40,232,534.42	951	2.36 (2.22–2.52)
56–65	2,296,730	10,703,243.75	3,491	32.62 (31.54–33.72)
66–75	1,353,725	6,069,656.00	14,132	232.83 (229.01–236.70)
76–85	814,111	3,735,595.68	42,120	1,127.53 (1,116.79–1,138.35)
≥86	307,040	1,021,646.46	24,127	2,361.58 (2,331.87–2,391.57)

InGef RDB=InGef Research Database, AD=Alzheimer’s disease, PYs=Person-years, CI=Confidence Interval, NA=Not applicable

Table S4. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in IQVIA DA Germany stratified by calendar year, sex, and age categories.

	IQVIA DA Germany			
	Number of participants	**Follow-up (PYs)**	**Number of individuals with AD**	**Incidence Rates/100,000 PYs (95% CI)**
Calendar year				
2014	7,414,784	6,063,812.40	2,517	41.51 (39.90–43.16)
2015	7,893,441	6,471,328.04	2,548	39.37 (37.86–40.93)
2016	8,567,236	6,890,132.73	3,038	44.09 (42.54–45.69)
2017	9,107,885	7,363,147.67	3,142	42.67 (41.19–44.19)
2018	9,324,436	7,581,357.56	3,498	46.14 (44.62–47.69)
2019	9,832,681	7,767,309.07	3,985	51.30 (49.72–52.92)
2020	10,183,708	7,997,904.30	3,990	49.89 (48.35–51.46)
2021	11,193,410	8,183,625.59	4,121	50.36 (48.83–51.92)
2022	10,836,728	7,973,288.69	4,895	61.39 (59.69–63.14)
2023	10,039,599	7,095,758.84	5,410	76.24 (74.22–78.30)
2024	8,660,179	5,296,238.12	5,230	98.75 (96.09–101.46)
Sex				
Female	16,767,231	45,455,927.21	25,150	55.33 (54.65–56.02)
Male	13,211,467	33,227,975.80	17,224	51.84 (51.06–52.62)
Age categories				
Early-onset AD (18–65)	22,731,283	54,244,486.36	2,992	5.52 (5.32–5.72)
Late-onset AD (≥66)	8,559,044	24,439,416.65	39,382	161.14 (159.55–162.74)
18–55	17,640,556	38,578,439.13	885	2.29 (2.14–2.45)
56–65	6,640,065	15,666,047.22	2,107	13.45 (12.88–14.04)
66–75	5,095,455	12,567,465.96	7,404	58.91 (57.58–60.27)
76–85	3,890,070	9,744,285.28	21,947	225.23 (222.26–228.23)
≥86	1,180,924	2,127,665.41	10,031	471.46 (462.27–480.77)

IQVIA DA Germany=IQVIA Disease Analyzer Germany, AD=Alzheimer's disease, PYs=Person-years, CI=Confidence Interval

Table S5. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in IPCI stratified by calendar year, sex, and age categories.

	IPCI			
	Number of participants	**Follow-up (PYs)**	**Number of individuals with AD**	**Incidence Rates/100,000 PYs (95% CI)**
Calendar year				
2014	852,075	739,723.27	488	65.97 (60.25–72.09)
2015	965,383	862,784.04	542	62.82 (57.64–68.34)
2016	1,000,439	939,236.53	640	68.14 (62.96–73.63)
2017	1,041,936	967,159.23	751	77.65 (72.19–83.41)
2018	1,039,654	977,833.29	744	76.09 (70.72–81.76)
2019	1,110,195	1,009,117.77	721	71.45 (66.33–76.86)
2020	1,190,342	1,084,985.45	686	63.23 (58.58–68.14)
2021	1,241,733	1,125,490.04	748	66.46 (61.78–71.40)
2022	1,233,603	1,034,060.45	646	62.47 (57.75–67.48)
2023	1,112,873	1,032,238.49	587	56.87 (52.36–61.66)
2024	1,146,392	1,066,078.61	487	45.68 (41.71–49.92)
Sex				
Female	1,089,621	5,571,052.73	4,374	78.51 (76.20–80.88)
Male	1,018,634	5,267,654.43	2,666	50.61 (48.71–52.57)
Age categories				
Early-onset AD (18–65)	1,751,541	8,382,516.46	311	3.71 (3.31–4.15)
Late-onset AD (≥66)	516,450	2,456,190.70	6,729	273.96 (267.45–280.59)
18–55	1,475,444	6,571,483.22	54	0.82 (0.62–1.07)
56–65	470,966	1,811,033.24	257	14.19 (12.51–16.04)
66–75	369,377	1,443,219.18	1,432	99.22 (94.15–104.50)
76–85	226,148	784,283.36	3,374	430.20 (415.81–444.97)
≥86	81,195	228,688.16	1,923	840.88 (803.72–879.33)

IPCI=Integrated Primary Care Information, AD=Alzheimer’s disease, PYs=Person-years, CI=Confidence Interval

Table S6. Incidence rates per 100,000 PYs (95% CI) of AD in the general adult population in CPRD GOLD stratified by calendar year, sex, and age categories.

	CPRD GOLD			
	Number of participants	Follow-up (PYs)	Number of individuals with AD	Incidence Rates/100,000 PYs (95% CI)
Calendar year				
2014	5,357,917	4,806,013.90	4,294	89.35 (86.69–92.06)
2015	4,858,512	4,313,870.76	4,125	95.62 (92.73–98.58)
2016	4,164,928	3,728,244.26	3,470	93.07 (90.00–96.22)
2017	3,752,735	3,409,632.09	3,194	93.68 (90.46–96.98)
2018	3,508,693	3,235,867.50	3,224	99.63 (96.22–103.13)
2019	3,396,074	3,117,778.75	3,365	107.93 (104.31–111.64)
2020	3,188,648	2,902,730.07	2,199	75.76 (72.62–78.99)
2021	2,929,958	2,618,412.50	2,756	105.25 (101.36–109.26)
2022	2,677,622	2,450,655.16	2,923	119.27 (114.99–123.68)
2023	2,554,253	2,374,057.49	2,843	119.75 (115.39–124.24)
2024	2,505,867	2,283,687.60	2,744	120.16 (115.70–124.74)
Sex				
Female	3,905,910	17,889,139.75	22,352	124.95 (123.32–126.60)
Male	3,698,560	17,351,810.33	12,785	73.68 (72.41–74.97)
Age categories				
Early-onset AD (18–65)	6,273,690	27,635,870.74	1,239	4.48 (4.24–4.74)
Late-onset AD (≥66)	1,755,787	7,605,079.34	33,898	445.73 (441.00–450.50)
18–55	5,366,545	22,219,192.62	218	0.98 (0.86–1.12)
56–65	1,455,717	5,416,678.12	1,021	18.85 (17.71–20.04)
66–75	1,131,660	4,251,164.74	5,691	133.87 (130.41–137.39)
76–85	730,207	2,449,082.10	16,655	680.05 (669.76–690.46)
≥86	339,080	904,832.50	11,552	1,276.70 (1,253.52–1,300.20)

CPRD GOLD=Clinical Practice Research Datalink GOLD, AD=Alzheimer's disease, PYs=Person-years, CI=Confidence Interval

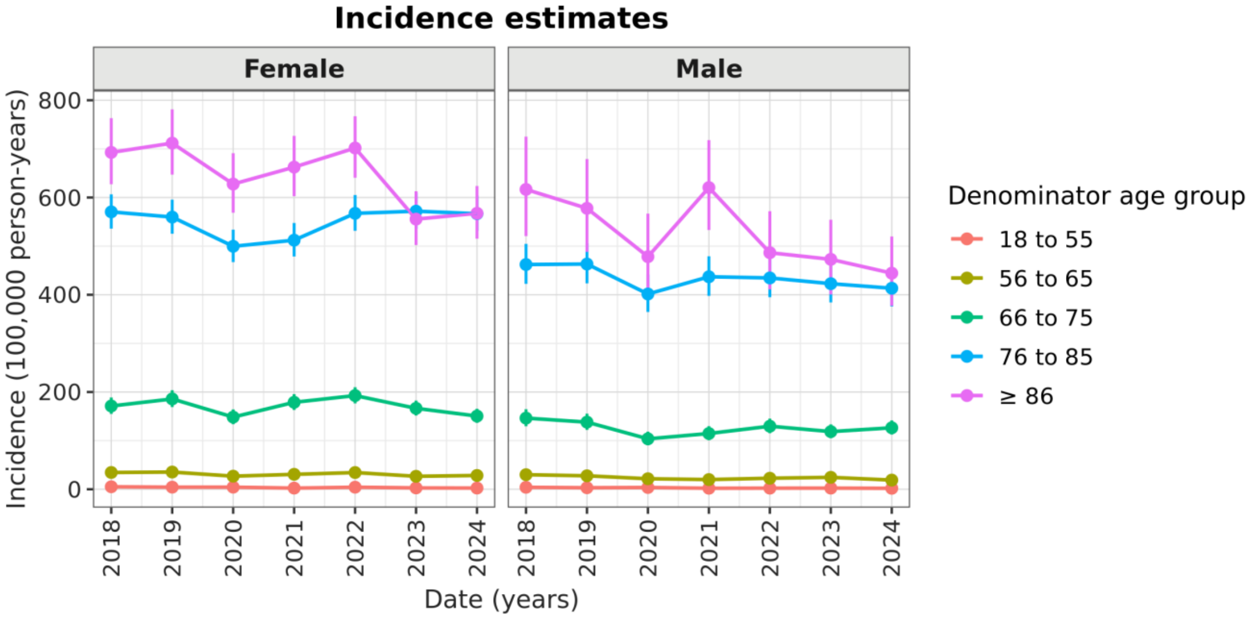

Figure S1. Yearly incidence rates per 100,000 PYs of AD in NAJS, stratified by sex and age categories.

NAJS=National Public Health Information System

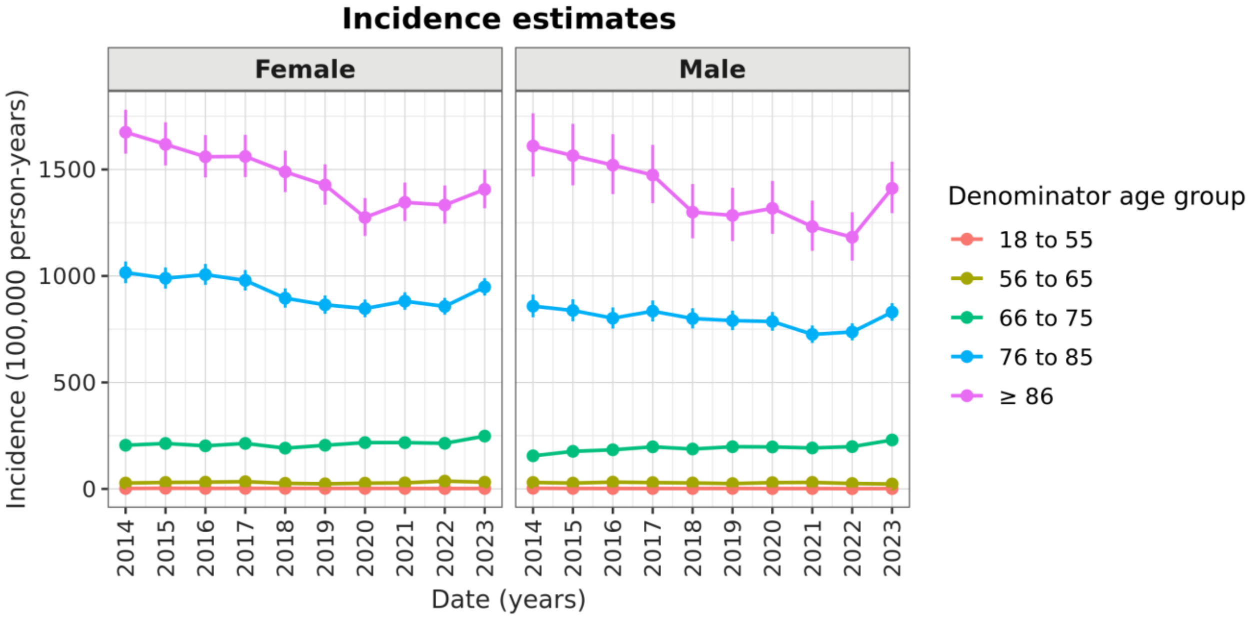

Figure S2. Yearly incidence rates per 100,000 PYs of AD in DK-DHR, stratified by sex and age categories.

DK-DHR=Danish Data Health Registries

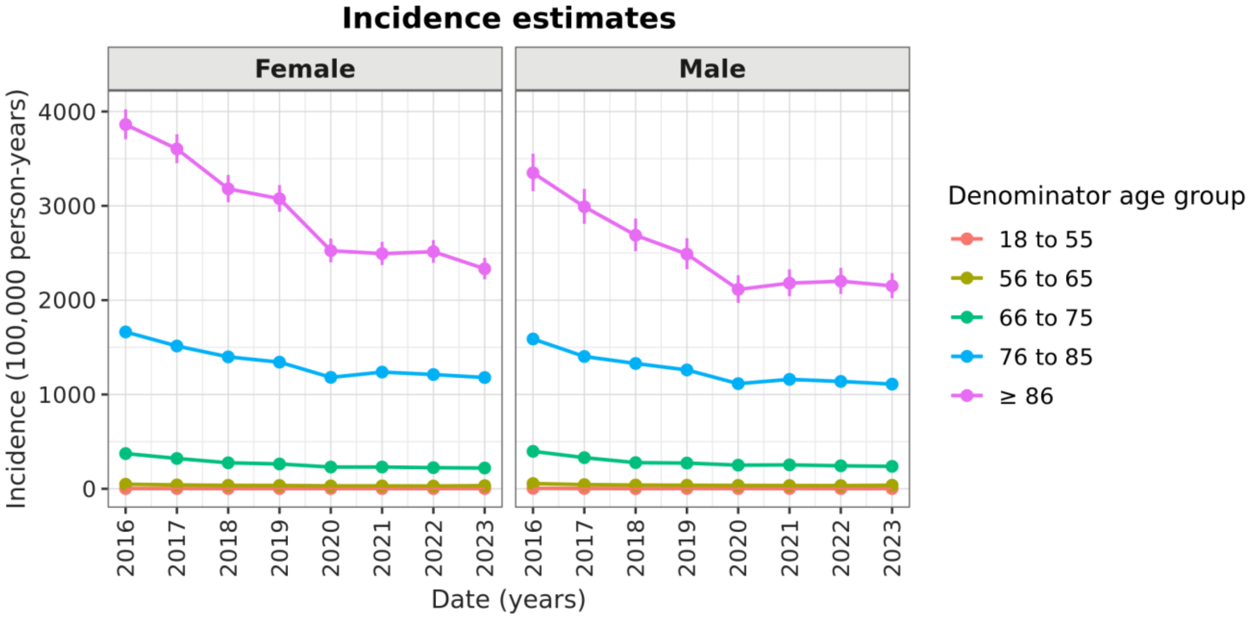

Figure S3. Yearly incidence rates per 100,000 PYs of AD in InGef RDB, stratified by sex and age categories.

InGef RDB=InGef Research Database

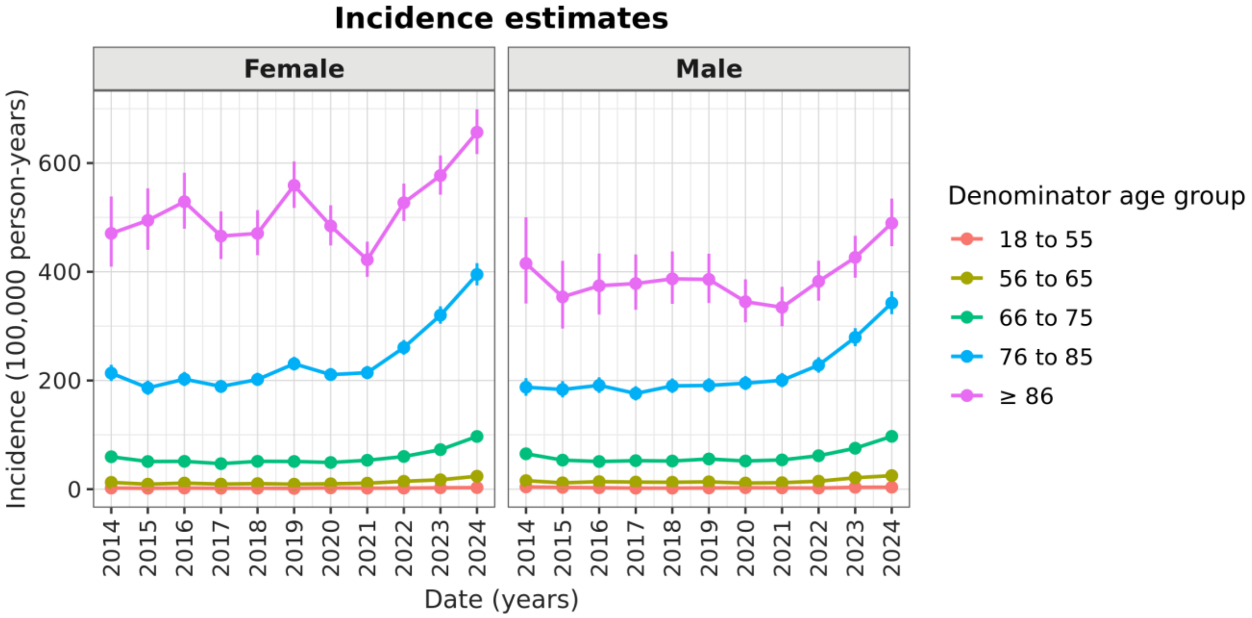

Figure S4. Yearly incidence rates per 100,000 PYs of AD in IQVIA DA Germany, stratified by sex and age categories.

IQVIA DA Germany=IQVIA Disease Analyzer Germany

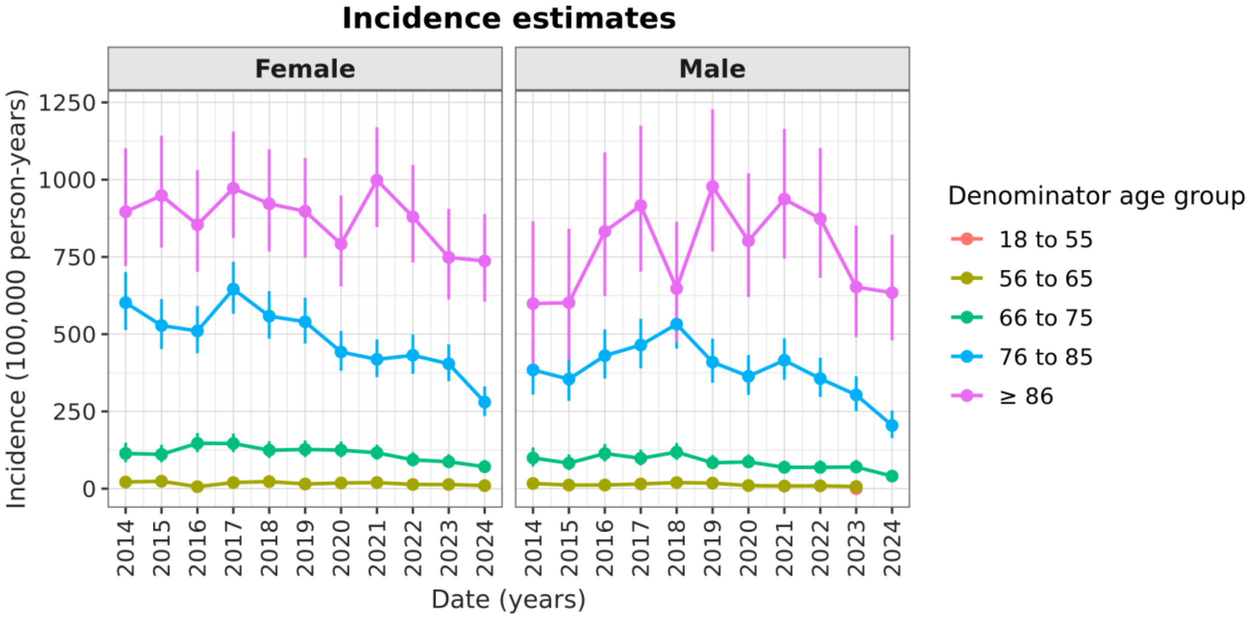

Figure S5. Yearly incidence rates per 100,000 PYs of AD in IPCI, stratified by sex and age categories.

IPCI=Integrated Primary Care Information

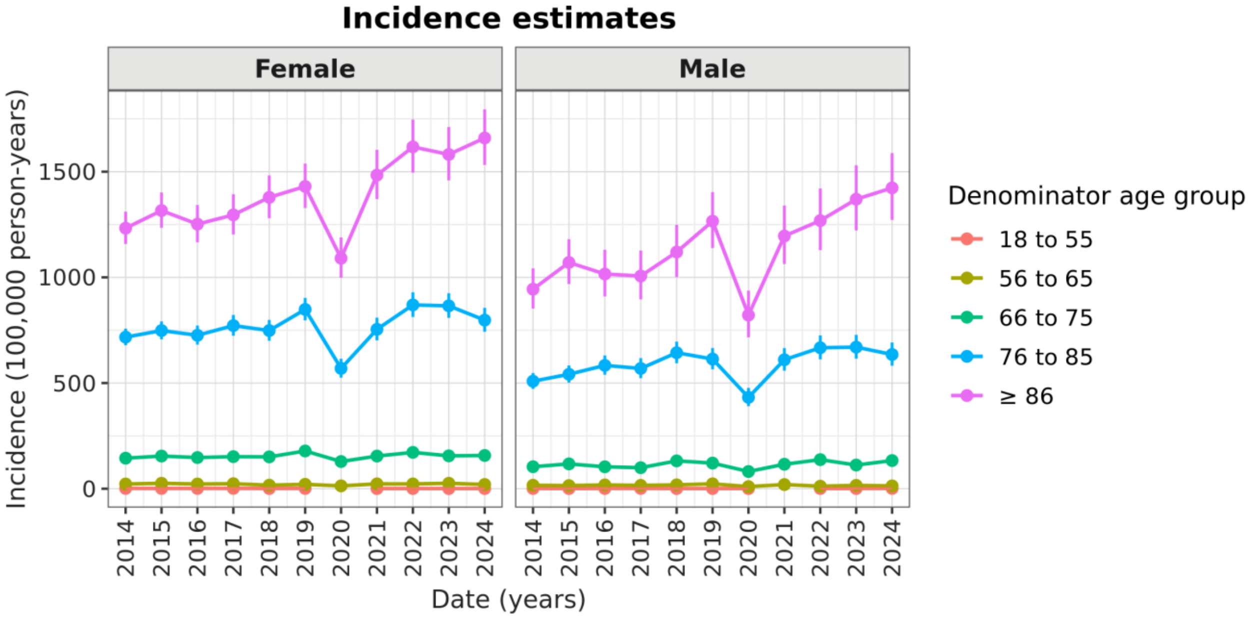

Figure S6. Yearly incidence rates per 100,000 PYs of AD in CPRD GOLD, stratified by sex and age categories.

CPRD=Clinical Practice Research Datalink

ANNEX VII: Prevalence of AD in the general adult population

Table S1. Prevalence (95% CI) of AD in the general adult population in NAJS stratified by calendar year, sex, and age categories.

	NAJS		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	NA	NA	NA
2015	NA	NA	NA
2016	NA	NA	NA
2017	NA	NA	NA
2018	3,481,350	8,254	0.24 (0.23–0.24)
2019	3,514,745	9,813	0.28 (0.27–0.28)
2020	3,538,129	10,755	0.30 (0.30–0.31)
2021	3,578,618	11,455	0.32 (0.31–0.33)
2022	3,585,517	12,248	0.34 (0.34–0.35)
2023	3,598,145	12,787	0.36 (0.35–0.36)
2024	3,620,631	13,481	0.37 (0.37–0.38)
Sex			
Female	2,124,165	18,443	0.87 (0.86–0.88)
Male	2,019,081	9,352	0.46 (0.45–0.47)
Age groups			
Early-onset AD (18–65)	3,324,110	2,260	0.07 (0.06–0.07)
Late-onset AD (≥66)	1,390,984	26,376	1.90 (1.87–1.92)
18–55	2,690,561	636	0.02 (0.02–0.03)
56–65	1,223,622	1,833	0.15 (0.14–0.16)
66–75	1,013,300	7,691	0.76 (0.74–0.78)
76–85	672,314	16,081	2.39 (2.36–2.43)
≥86	262,292	7,959	3.03 (2.97–3.10)

NAJS=National Public Health Information System, AD=Alzheimer's disease, CI=Confidence Interval, NA=Not applicable

Table S2. Prevalence (95% CI) of AD in the general adult population in DK-DHR stratified by calendar year, sex, and age categories.

	DK-DHR		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	4,572,916	28,224	0.62 (0.61–0.62)
2015	4,621,939	29,416	0.64 (0.63–0.64)
2016	4,664,299	30,217	0.65 (0.64–0.66)
2017	4,700,233	31,045	0.66 (0.65–0.67)
2018	4,734,810	31,500	0.66 (0.66–0.67)
2019	4,761,092	31,720	0.67 (0.66–0.67)
2020	4,776,748	32,259	0.68 (0.67–0.68)
2021	4,807,446	32,735	0.68 (0.67–0.69)
2022	4,874,519	32,970	0.68 (0.67–0.68)
2023	4,914,666	34,021	0.69 (0.69–0.70)
2024	NA	NA	NA
Sex			
Female	3,349,345	67,357	2.01 (2.00–2.03)
Male	3,315,170	43,380	1.31 (1.30–1.32)
Age groups			
Early-onset AD (18–65)	5,896,736	5,494	0.09 (0.09–0.10)
Late-onset AD (≥66)	2,121,970	108,046	5.09 (5.06–5.12)
18–55	5,199,785	1,291	0.03 (0.02–0.03)
56–65	2,277,678	5,030	0.22 (0.22–0.23)
66–75	1,773,391	25,908	1.46 (1.44–1.48)
76–85	1,188,270	69,295	5.83 (5.79–5.87)
≥86	468,488	48,606	10.38 (10.29–10.46)

DK-DHR=Danish Data Health Registries, AD=Alzheimer’s disease, CI=Confidence Interval, NA=Not applicable

Table S3. Prevalence (95% CI) of AD in the general adult population in InGef RDB stratified by calendar year, sex, and age categories.

	InGef RDB		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	NA	NA	NA
2015	NA	NA	NA
2016	6,365,230	12,080	0.19 (0.19–0.19)
2017	6,386,106	21,515	0.34 (0.33–0.34)
2018	6,423,040	28,430	0.44 (0.44–0.45)
2019	6,466,292	33,746	0.52 (0.52–0.53)
2020	6,461,248	37,224	0.58 (0.57–0.58)
2021	6,514,161	39,897	0.61 (0.61–0.62)
2022	6,565,036	42,278	0.64 (0.64–0.65)
2023	6,630,010	43,504	0.66 (0.65–0.66)
2024	NA	NA	NA
Sex			
Female	4,365,865	47,633	1.09 (1.08–1.10)
Male	4,470,762	37,189	0.83 (0.82–0.84)
Age groups			
Early-onset AD (18–65)	7,813,829	4,443	0.06 (0.06–0.06)
Late-onset AD (≥66)	1,802,627	81,935	4.54 (4.52–4.58)
18–55	6,817,836	952	0.01 (0.01–0.01)
56–65	2,297,199	3,960	0.17 (0.17–0.18)
66–75	1,355,281	15,688	1.16 (1.14–1.18)
76–85	820,534	48,542	5.92 (5.86–5.97)
≥86	319,352	36,437	11.41 (11.30–11.52)

InGef RDB=InGef Research Database, AD=Alzheimer's disease, CI=Confidence Interval, NA=Not applicable

Table S4. Prevalence (95% CI) of AD in the general adult population in IQVIA DA Germany stratified by calendar year, sex, and age categories.

	IQVIA DA Germany		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	7,420,573	8,303	0.11 (0.11–0.11)
2015	7,900,484	9,591	0.12 (0.12–0.12)
2016	8,575,274	11,075	0.13 (0.13–0.13)
2017	9,117,170	12,421	0.14 (0.13–0.14)
2018	9,334,769	13,828	0.15 (0.15–0.15)
2019	9,844,088	15,383	0.16 (0.15–0.16)
2020	10,196,211	16,483	0.16 (0.16–0.16)
2021	11,206,556	17,258	0.15 (0.15–0.16)
2022	10,850,204	18,363	0.17 (0.17–0.17)
2023	10,053,698	19,495	0.19 (0.19–0.20)
2024	8,674,675	19,719	0.23 (0.22–0.23)
Sex			
Female	16,770,623	28,540	0.17 (0.17–0.17)
Male	13,213,864	19,620	0.15 (0.15–0.15)
Age groups			
Early-onset AD (18–65)	22,731,857	3,566	0.02 (0.01–0.02)
Late-onset AD (≥66)	8,565,191	45,526	0.53 (0.53–0.54)
18–55	17,640,760	1,089	0.01 (0.01–0.01)
56–65	6,640,827	2,869	0.04 (0.04–0.04)
66–75	5,097,515	9,464	0.19 (0.18–0.19)
76–85	3,896,527	28,400	0.73 (0.72–0.74)
≥86	1,187,819	16,923	1.43 (1.40–1.45)

IQVIA DA Germany=IQVIA Disease Analyzer Germany, AD=Alzheimer’s disease, CI=Confidence Interval

Table S5. Prevalence (95% CI) of AD in the general adult population in IPCI stratified by calendar year, sex, and age categories.

	IPCI		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	852,884	1,297	0.15 (0.14–0.16)
2015	966,301	1,460	0.15 (0.14–0.16)
2016	1,001,599	1,800	0.18 (0.17–0.19)
2017	1,043,379	2,194	0.21 (0.20–0.22)
2018	1,041,367	2,457	0.24 (0.23–0.24)
2019	1,112,151	2,677	0.24 (0.23–0.25)
2020	1,192,436	2,780	0.23 (0.22–0.24)
2021	1,243,854	2,869	0.23 (0.22–.24)
2022	1,235,708	2,751	0.22 (0.21–0.23)
2023	1,114,713	2,427	0.22 (0.21–0.23)
2024	1,148,147	2,242	0.20 (0.19–0.20)
Sex			
Female	1,090,152	4,905	0.45 (0.44–0.46)
Male	1,018,912	2,944	0.29 (0.28–0.30)
Age groups			
Early-onset AD (18–65)	1,751,599	369	0.02 (0.02–0.02)
Late-onset AD (≥66)	517,347	7,626	1.47 (1.44–1.51)
18–55	1,475,456	66	0.00 (0.00–0.01)
56–65	471,047	338	0.07 (0.06–0.08)
66–75	369,695	1,750	0.47 (0.45–0.50)
76–85	227,221	4,447	1.96 (1.90–2.02)
≥86	82,240	2,968	3.61 (3.48–3.74)

IPCI=Integrated Primary Care Information, AD=Alzheimer’s disease, CI=Confidence Interval

Table S6. Prevalence (95% CI) of AD in the general adult population in CPRD GOLD stratified by calendar year, sex, and age categories.

	CPRD GOLD		
	Number of individuals with AD	**Number of participants**	**Prevalence (95% CI)**
Calendar year			
2014	5,370,677	17,054	0.32 (0.31–0.32)
2015	4,871,438	17,051	0.35 (0.34–0.36)
2016	4,176,890	15,432	0.37 (0.36–0.38)
2017	3,763,675	14,134	0.38 (0.37–0.38)
2018	3,519,122	13,653	0.39 (0.38–0.40)
2019	3,406,578	13,869	0.41 (0.40–0.41)
2020	3,199,143	12,694	0.40 (0.39–0.40)
2021	2,939,049	11,847	0.40 (0.40–0.41)
2022	2,686,356	11,657	0.43 (0.43–0.44)
2023	2,562,893	11,483	0.45 (0.44–0.46)
2024	2,514,740	11,617	0.46 (0.45–0.47)
Sex			
Female	3,914,373	30,815	0.79 (0.78–0.80)
Male	3,702,857	17,082	0.46 (0.45–0.47)
Age groups			
Early-onset AD (18–65)	6,274,079	1,628	0.03 (0.03–0.03)
Late-onset AD (≥66)	1,768,736	46,847	2.65 (2.62–2.67)
18–55	5,366,590	263	0.00 (0.00–0.00)
56–65	1,456,170	1,474	0.00 (0.00–0.00)
66–75	1,134,073	8,104	0.01 (0.01–0.01)
76–85	738,958	25,406	0.03 (0.03–0.03)
≥86	349,360	21,832	0.06 (0.06–0.06)

CPRD GOLD=Clinical Practice Research Datalink GOLD, AD=Alzheimer’s disease, CI=Confidence Interval

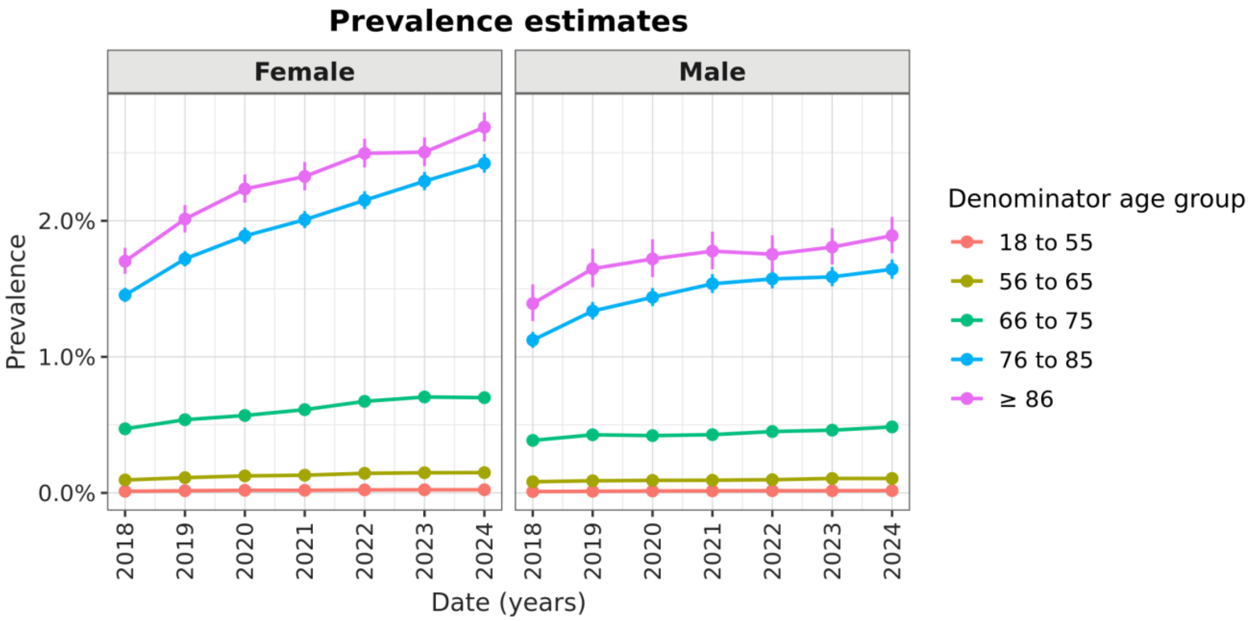

Figure S1. Prevalence (95% CI) of AD in NAJS, stratified by sex and age categories.

NAJS=National Public Health Information System

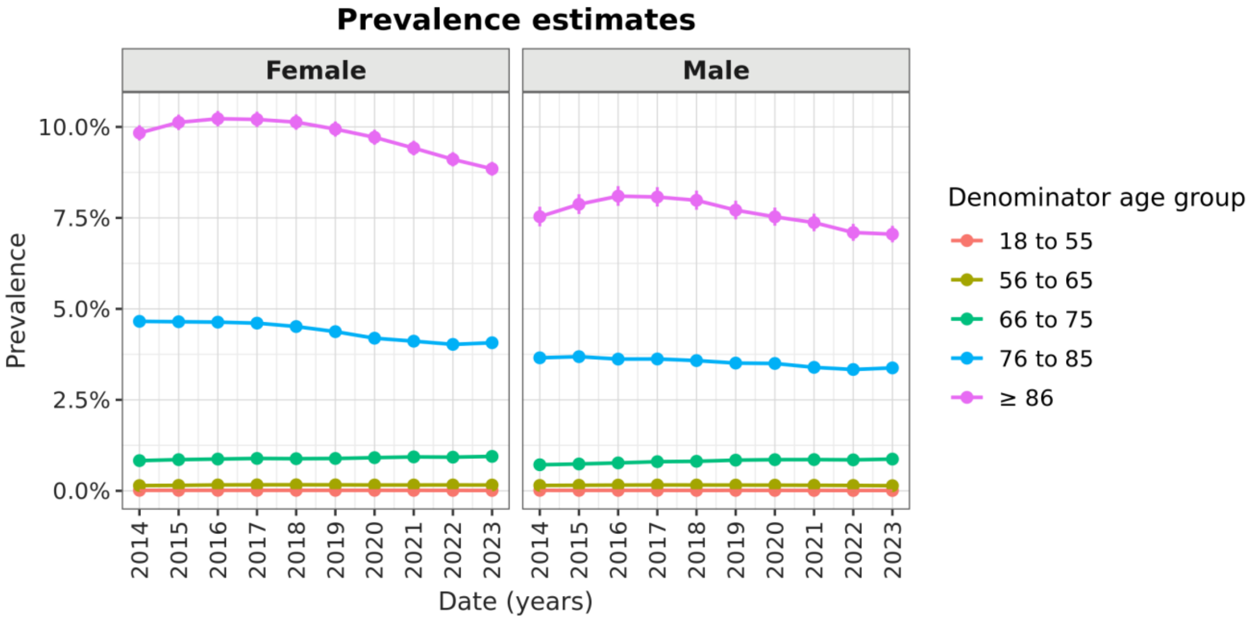

Figure S2. Prevalence (95% CI) of AD in DK-DHR, stratified by sex and age categories.

DK-DHR=Danish Data Health Registries

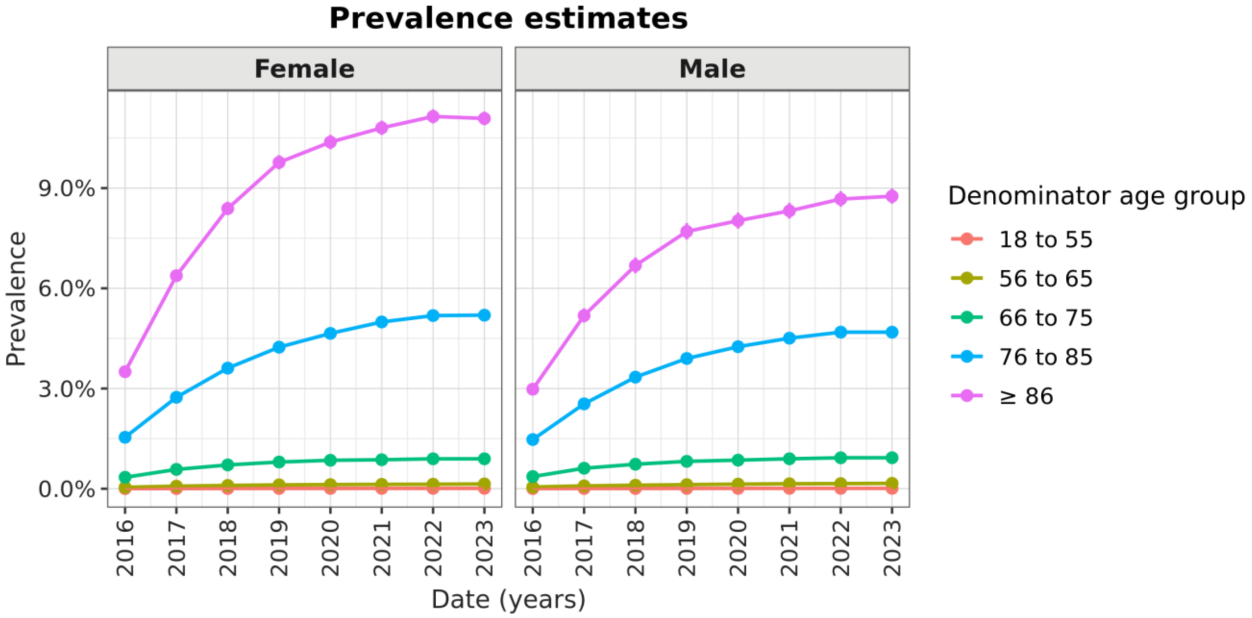

Figure S3. Prevalence (95% CI) of AD in InGef RDB, stratified by sex and age categories.

InGef RDB=InGef Research Database

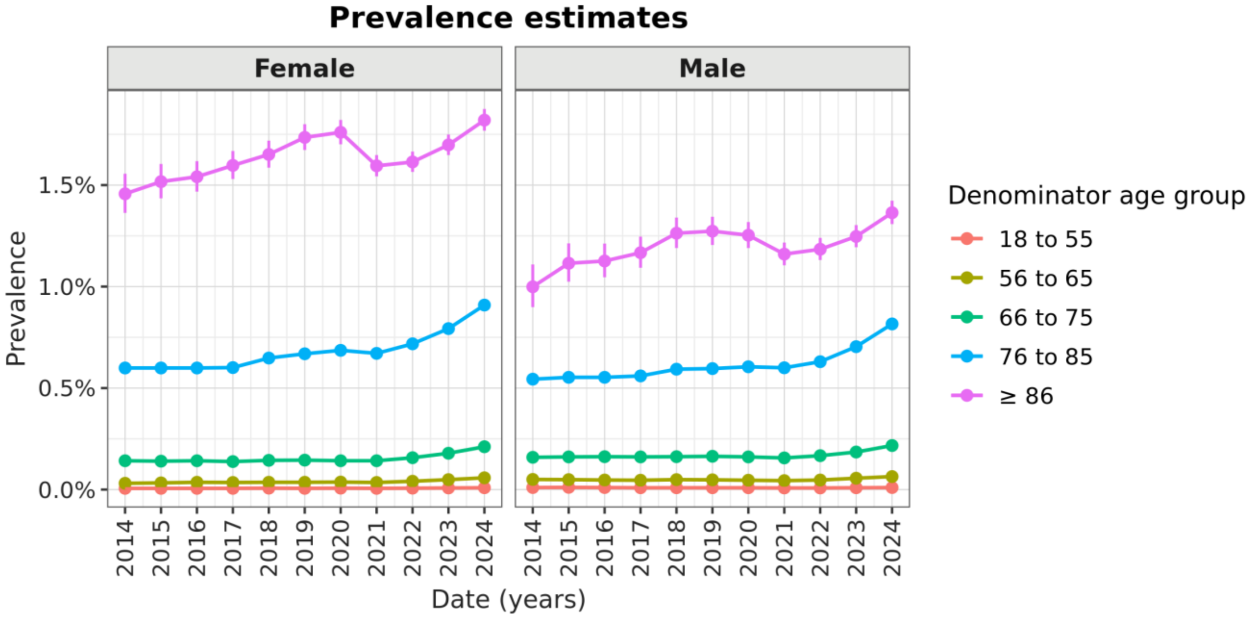

Figure S4. Prevalence (95% CI) of AD in IQVIA DA Germany, stratified by sex and age categories.

IQVIA DA Germany=IQVIA Disease Analyzer Germany

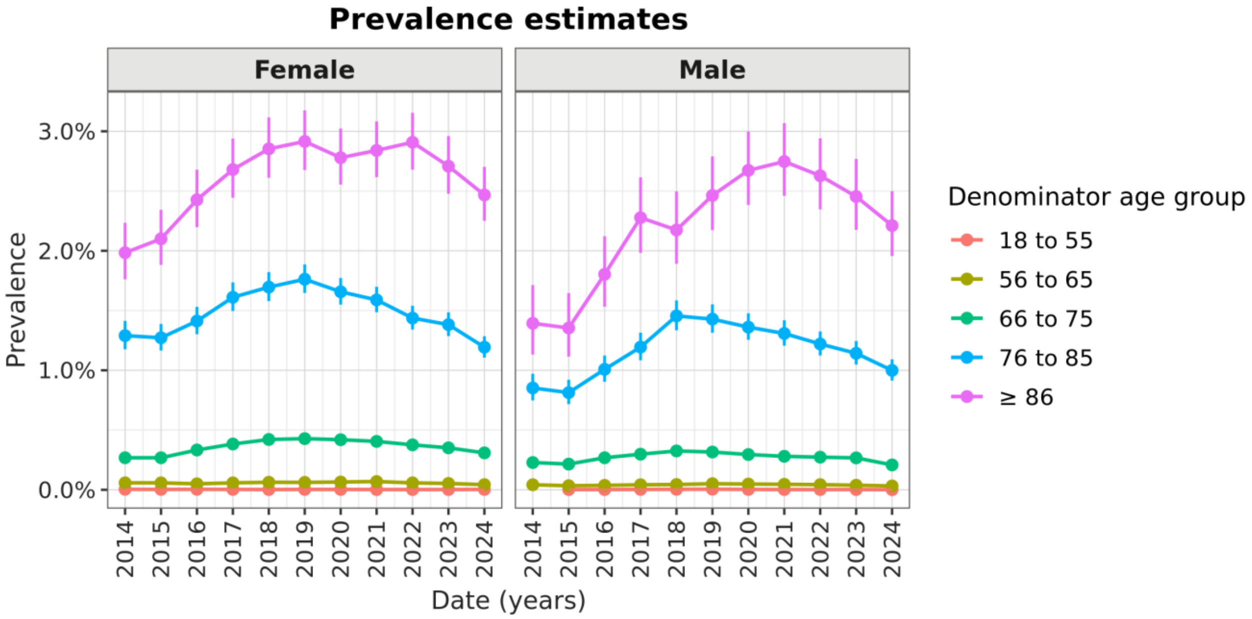

Figure S5. Prevalence (95% CI) of AD in IPCI, stratified by sex and age categories.

IPCI=Integrated Primary Care Information

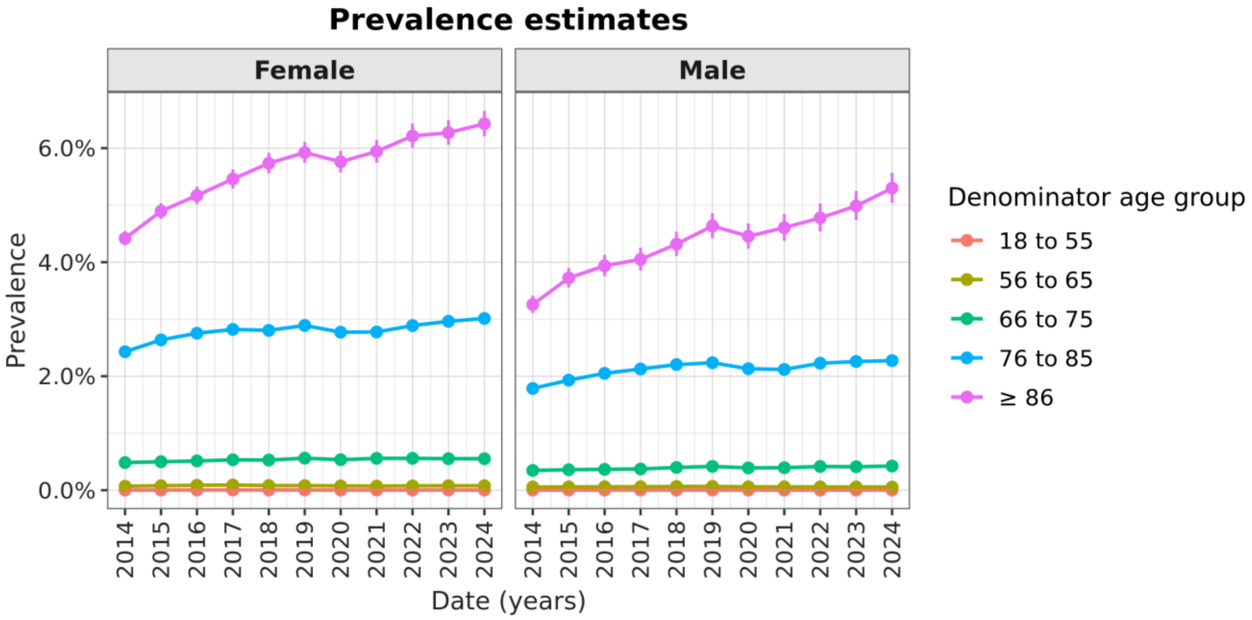

Figure S6. Prevalence (95% CI) of AD in CPRD GOLD, stratified by sex and age categories.

CPRD=Clinical Practice Research Datalink

ANNEX VIII: Glossary

Additional definitions are available in the EMA Glossary of terms https://www.ema.europa.eu/en/about-us/glossaries.

Aggregated Data

Data collected and combined from multiple sources to generate summary information, typically anonymised.

Benefit-Risk Assessment

Evaluation of the positive therapeutic effects of a medicine compared to its risks (e.g., side effects).

Common Data Model (CDM)

A standardized data structure that enables data from multiple sources to be harmonized, making analysis consistent and reproducible. DARWIN EU® utilises the OMOP CDM maintained by the OHDSI community.

Complex Studies (C3)

Studies requiring the development or customisation of specific study designs, protocols, and Statistical Analysis Plans (SAPs), with extensive collection or extraction of data. Examples include etiological studies measuring the strength and determinants of an association between an exposure and the occurrence of a health outcome in a defined population considering sources of bias, potential confounding factors, and effect modifiers.

Coordination Centre (CC)

The central hub responsible for managing and overseeing the activities within DARWIN EU®. It is based at Erasmus University Medical Centre in Rotterdam, the Netherlands.

Data Access

The process of obtaining permission to use specific datasets for regulatory or scientific studies.

Data Quality Framework

A set of standards and procedures to ensure accuracy, completeness, timeliness, and consistency of data used in DARWIN EU®.

Data Source

A database or repository of structured health-related data, such as electronic health records (EHRs), insurance claims, or registries.

DARWIN EU®

The European Medicines Agency's (EMA) federated network of real-world data sources designed to generate evidence to support regulatory decision-making.

EMA (European Medicines Agency)

The regulatory body responsible for the evaluation and supervision of medicinal products in the EU, overseeing DARWIN EU®.

Evidence Generation

The process of analysing real-world data to produce scientific information that can inform healthcare or regulatory decisions.

Federated Network

A data infrastructure where data remain at their original location but can be analysed in a harmonised way across multiple partners using a common model and tools.

GDPR (General Data Protection Regulation)

The EU regulation governing the protection of personal data and privacy, crucial to how DARWIN EU® handles health data.

Health Technology Assessment (HTA)

A systematic evaluation of properties and impacts of health technology, often using DARWIN EU® data to support assessments.

Metadata

Descriptive information about a data source (e.g., its content, quality, and structure), essential for identifying relevant databases in DARWIN EU® studies.

Off-the-Shelf Studies (OTS)

Studies for which a standard protocol per study/analysis type and standardised analytics may be developed and applied or adapted, typically relating to a descriptive research question. This includes studies on disease epidemiology, for example, the estimation of the prevalence or incidence of health outcomes in defined time periods and population groups, or drug utilisation studies at the population or patient level.

OHDSI (Observational Health Data Sciences and Informatics)

An open-science collaborative community that develops tools and standards (including the OMOP CDM) to enable large-scale analytics of observational health data. OHDSI provides the technical and scientific foundation for DARWIN EU®'s analytical ecosystem.

Patient-Level Data

Data related to individuals, de-identified, used for longitudinal or detailed analyses.

OMOP (Observational Medical Outcomes Partnership)

A common data model (CDM) that standardises the structure and content of observational healthcare data, enabling systematic analysis across disparate datasets. DARWIN EU® uses the OMOP CDM to ensure interoperability and consistency in real-world evidence generation.

Real-World Data (RWD)

Data relating to individual health status or healthcare delivery that is collected from routine clinical practice rather than from randomised controlled trials.

Real-World Evidence (RWE)

Clinical evidence derived from the analysis of RWD, used to inform decisions by regulators, payers, or clinicians.

Regulatory Decision-Making

The process by which authorities like EMA assess data to authorise, monitor, or modify the use of medicines in the EU.

Routine Repeated Studies (RR)

Studies that are either Off-the-Shelf or Complex studies repeated on a regular basis, following the same protocol and study code, but with updated data and/or different data partners.

Study Protocol

A detailed plan describing how a specific real-world study will be conducted, including objectives, design, data sources, and analyses.

Very Complex Studies (C4)

Studies which cannot rely only on electronic health care databases, or which would require complex methodological work, for example, due to the occurrence of events that cannot be defined by existing diagnosis codes, including events that do not yet have a diagnosis code, where it may be necessary to combine a diagnosis code with other data such as results of laboratory investigations. These studies might require the collection of data prospectively, or the inclusion of new (not previously onboarded) data sources.