Universität Rostock Traditio et Innovatio

Dementia Risk Factors Over the Life Course – Identification, Trajectories and Pathways

Dissertation

zur Erlangung des akademischen Grades

doctorum rerum politicarum (Dr. rer. pol)

am Institut für Soziologie und Demographie

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von

Constantin Reinke, geb. am 10.01.1991 in Hamburg

Wohnhaft in Rostock

Rostock, den 02.12.2024

https://doi.org/10.18453/rosdok_id00005102

GutachterInnen:

Prof. Dr. Gabriele Doblhammer, Universität Rostock, Institut für Soziologie und Demographie

Prof. Dr. Matthias Schmid, Universität Bonn, Institut für Medizinische Biometrie, Informatik und Epidemiologie (IMBIE)

Jahr der Einreichung: 2024

Jahr der Verteidigung: 2025

ACKNOWLEDGEMENTS

This work has been characterized by great personal experiences and challenges, the discovery of novel paths and opportunities, and most of all by new insights. I have met many people from whom I have learned a lot and without whom I probably would not have come this far or even travelled this path. I would like to thank all the people who supported me in or during this work.

First of all, I would like to thank my supervisor and mentor Gabriele Doblhammer. Gabriele, you made it possible for me to do this work in the first place. You always supported and encouraged me and believed in me. Becoming part of your team at the University of Rostock enabled me to take part in scientific exchange, many conferences, workshops, and to gain insights into projects beyond my thesis. These experiences have shaped me both professionally and personally.

I would also like to thank my co-supervisor Matthias Schmid. Matthias, you gave me an introduction to the field of statistical learning. You also enabled me to do a research stay at your institute, the IMBIE in Bonn, which greatly enriched my methodological knowledge and was a great experience for me.

Furthermore, I wish to express my gratitude to my co-authors: Katharina Heser, Michael Wagner, Nikolaus Buchmann, Christina Tegeler-Nonnen, Ilja Demuth, and Thomas Welchowski. Thank you for the good and successful collaboration.

A special thanks to my co-author and colleague Anne Fink: Anne, not only have I learned a lot about health claims data from you, but I have always been able to rely on your experience; you are an important mentor to me. I am very grateful to you for that.

I would also like to thank all my recent and former colleagues for the great teamwork and the constructive and helpful feedback to my work: Benjamin Aretz, Poulami Barman, Alexander Barth, Maximilian Frentz-Göllnitz, Thomas Fritze, Daniela Georges, Anna-Viktoria Holtz, Daniel Kreft, Michael Nerius, Sophia Nestler, Svenja Östreich, Elena Rakuša, Roland Rau, Niklas Ullrich-Kniffka, Bärbel Westendorf, and Christina Westphal.

I am grateful to Rafael Weißbach, who gave me a lot of responsibility during my time as a research assistant and thus provided me with valuable insights into the processes at the university and in academia. These experiences strengthened my decision to pursue a doctorate.

Thanks to my friends and family for their emotional support along the way. Finally, I would like to thank Nina for her unshakeable support all along the way. I probably wouldn't have made it this way without you.

I am grateful to the Scientific Research Institute of the AOK, WIdO, for providing the data used for the underlying Studies 1, 3 & 4 of this thesis.

The services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participants are gratefully acknowledged. The Lifelines initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG), Groningen University and the Provinces in the North of the Netherlands (Drenthe, Friesland, Groningen).

SUMMARY

Currently between 55-57 million individuals world-wide are affected by dementia; a number expected to reach up to 152 million by the year 2050. Dementia is a severe burden for those affected, their caregivers and the healthcare systems. Although a cure for the disease may still seem out of reach, the progression of the disease can be delayed if interventions start at earlier stages of the disease, emphasizing the importance of early detection. However, the primary goal should be to prevent dementia. In order to achieve this, it is essential to identify and obtain a better understanding of the risk factors involved. The objective of this thesis was to identify risk factors and explore their relations to the trajectories of the risk of dementia. The results contribute to supporting the prevention of dementia and facilitating the early detection of such cases, thereby delaying the progression of the disease. The overarching objective of this thesis was addressed in four studies.

To define the theoretical background for the thesis, the outcome measures of dementia and cognitive function were introduced as well as the exposures of education, diabetes and late-life depression and their empirical connections to the outcomes.

The underlying studies were contextualized into the theoretical framework of life course epidemiology. This framework combines approaches from various fields related to population health sciences and considers the timing and duration effects of risk factors as well as their accumulation and interaction for health outcomes over the entire life course.

Two different data sources provided the basis for the analyses of the underlying studies. A sample of German health claims data (AOK) with 250,000 individuals with 13 years of follow-up were used in the Studies 1,3, and 4. The analyses in Study 2 were based on Lifelines, a prospective population-based cohort study and biobank of 167,729 individuals from the Netherlands

Study 1 addressed the questions whether German health claims data is suitable for dementia risk predictions and which important risk factors can be identified. Logistic regression and methods of machine learning (gradient boosting machines and random forest models) were applied to build prediction models on a large sample from the largest German health insurance company AOK. The evaluation of the prediction models revealed acceptable prediction accuracy and identified established as well as new predictors associated with the risk of dementia. Predictions from German health claims data may provide a decision support tool for early dementia screening in a cost-effective manner.

Study 2 aimed at exploring the association of education and cognitive function by questing if diabetes may act as a mediator on this pathway. Data from the large population-based health survey Lifelines from the Netherlands was used to apply regression models and mediation analyses. The results

suggested that a small part of the effect of education on cognitive function may act through diabetes. People with lower levels of education may be twice as disadvantaged in terms of cognitive function, indirectly through their higher risk of diabetes and directly through their education.

Study 3 examined the temporal trajectory of the risk of dementia in individuals diagnosed with diabetes as well as the potentially modifying effects of diabetes severity and treatment type on this association. The risk of dementia over time since first diagnosis of type 2 diabetes was U-shaped. Across the different diabetes treatment groups, the U-shaped pattern was consistent. The U-shape was more pronounced in less severe cases with no diabetes complications at baseline. As a conclusion, physicians should maintain cognitive monitoring in diabetic patients beyond two years post-diagnosis, considering treatment outcomes and treatment adherence.

To explore if late-life depression is a risk factor or an early symptom of dementia, the temporal association between incident late-life depression and the development of subsequent dementia was explored in Study 4. We modeled the trajectory of the dementia risk over time since incident late-life depression by exponential hazard models from German claims data. We found a higher risk of dementia in subsequence of incident late-life depression. The strongest association was found for the shortest interval between both diagnoses. Furthermore, the association was found to be particularly strong in men and for individuals younger than 74 years of age. It can be concluded that late-life depression can be a prodrome of dementia and signal an elevated risk of dementia in the near future.

The discussion of strengths and limitations of the underlying studies revealed important issues concerning the availability or accessibility of data. A lot of valuable information is existing in routinely collected data but still not accessible or linkable. Given the current limited evidence base for some potential factors, further research into modifiable risk factors is necessary and efforts should be made to improve the availability and accessibility of data. In addition to the new insights gained from this thesis and the underlying studies, which make a valuable contribution for the prevention and early detection of dementia, new questions emerged that should be addressed in future research. These include the transferability of the approaches into further contexts and with other outcomes.

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

95% CI	**95% Confidence Interval**
aDCSI	**Adopted Diabetes Complications Severity Index**
ADM	**Anti-Diabetic Medications**
AOK	**Allgemeine Ortskrankenkassen**
ATC	**Anatomical Therapeutic Chemical Classification**
BfArM	**Bundesinstitut für Arzneimittel und Medizinprodukte**
C-index	**Concordance Index**
EHR	**Electronic Health Records**
EMA	**European Medicines Agency**
GBM	**Gradient Boosting Machines**
HbA1c	**Glycated hemoglobin (A1c)**
HR	**Hazard Ratio**
ICD-10	**10th revision of the International Classification of Diseases**
LR	**Logistic Regression**
MCI	**Mild cognitive impairment**
ML	**Machine Learning**
mmol/L	**Millimole per liter**
OPS	**Operationen- und Prozedurenschlüssel – German adaptation of the International Classification of Procedures in Medicine**
OR	**Odds Ratio**
RF	**Random Forest**

LIST OF ORIGINAL PUBLICATIONS

Study 1

Reinke, C., Doblhammer, G., Schmid, M., Welchowski, T. (2022). Dementia risk predictions from German claims data using methods of machine learning. *Alzheimer's & Dementia*. 19 (2), 477-486.

Study 2

Reinke, C. (2024). The effect of diabetes in the multifaceted relationship between education and cognitive function. *BMC Public Health.* 24, 2584.

Study 3

Reinke, C., Buchmann, N., Fink, A., Tegeler, C., Demuth, I., Doblhammer, G. (2022). Diabetes duration and the risk of dementia: a cohort study based on German health claims data. *Age and Ageing*, 51 (1).

Study 4

Heser, K., Fink, A., **Reinke, C.**, Wagner, M., Doblhammer, G. (2020) The temporal association between incident late-life depression and incident dementia. *Acta psychiatrica Scandinavica*. 142 (5), 402-412.

Table of Contents

1. Introduction

During the last few years, which have been dominated by the Covid-19 pandemic, some important public health issues have fallen behind in the public discourse. One of the most important is dementia, which, despite alarming figures, is still a 'silent pandemic' in many areas. Although prevention is possible and, above all, urgently needed.

Dementia and normal brain ageing may both be accompanied by a deterioration of cognitive function (Wyss-Coray 2016; Denver and McClean 2018). However, only dementia is characterized by a significant impairment of an individual's daily functioning, which consequently leads to a reduction in quality of life and a loss of autonomy as well as a substantial burden on caregivers and healthcare systems. The decline in cognitive function can be conceptualized as a continuous process. The continuum between normal brain ageing and dementia is defined as mild cognitive impairment (MCI), which represents the pre-clinical phase of cognitive function decline. Dementia marks the point when the cognitive impairment reaches a clinical state (Pandya et al. 2016; Aisen et al. 2017). Typically, the process of decline in cognitive function begins many years before the manifestation of dementia (Bäckman et al. 2004).

According to the most recent estimates, there are approximately between 55-57 million individuals across the globe who are affected by dementia (World Health Organization 2023), which impairs their memory, thinking, and social abilities. Projections indicate that this figure is likely to increase drastically to around 78 million by the year 2030 (World Health Organization 2021), and is expected to reach up to 152 million by the year 2050 (Nichols et al. 2022), reflecting a concerning trend in the prevalence of dementia. The global social costs of dementia are about $1313.4 billion or $23,769 per person with dementia (Wimo et al. 2023). Dementia is already one of the most prevalent diseases at the time of death, and in this respect its importance will continue to increase in the coming decades (Doblhammer et al. 2022; World Health Organization 2024).

These alarming figures illustrate the urgent necessity for the fight against dementia to become a central focus of public health in the coming years and decades. Even if the challenges at political level are well known and the development and implementation of public health measures and programs has already been tackled (bmfsfj 2020; HHS 2023), there is an urgent need for more starting points and adjustments in order to be able to launch further programs and measures. In order to achieve this, it is essential to **identify and obtain a better understanding of the risk factors** involved. This is the role of demographic, epidemiological and public health research. In its last report on *Dementia prevention, intervention, and care* in 2020, the Lancet Commission came up with an estimated potential of 40% avoidable cases of dementia due to just 12 risk factors, based on the evidence available at the time

(Livingston et al. 2020). In the Lancet Commission's current report (Livingston et al. 2024), however, there are already 14 factors and the estimated potential for preventable cases of dementia has risen to 45%. This development clearly shows how important and useful further research into risk factors is in the fight against dementia. The factors that contribute to the theoretical proportion of preventable dementia cases are distributed across the entire life course.

A further point that serves to illustrate the importance of research into risk factors is provided by a consideration of the current status of dementia therapy. Although a cure for dementia remains beyond the current horizon of research, the progression of the disease can be effectively delayed, and the associated cognitive decline and symptoms can be reduced. The newest developments in pharmacological research have revealed promising progress in combating the abnormal degeneration of brain cells associated with dementia. The underlying mechanism involves the amyloid-ß plaques, which have emerged as a key target for therapeutic intervention (Avgerinos et al. 2021). The reports on the announcements of monoclonal antibody agents like Lecanemab (van Dyck et al. 2023) and Donanemab (Sims et al. 2023) by the pharmaceutical companies triggered a veritable euphoria, at least in the public perception, with regard to dementia research (The Lancet 2022). The research into the effectiveness of these agents is still relatively new and the amount of the clinical benefits is still being discussed (Høilund-Carlsen et al. 2024; Song et al. 2024). Nevertheless, the currently available evidence confirms that the amyloid antibodies reduce or rather remove amyloid-ß plaques by triggering an immune response, which can be seen as a decisive advance in the treatment of dementia (Klein et al. 2024; Gueorguieva et al. 2023; Qiao et al. 2024). While the agents have already been approved in the USA, the European Medicines Agency (EMA) also came to the conclusion on 14th November 2024 that benefits outweigh risks in a restricted patient population and recommended the marketing authorization of Lecanemab (EMA 2024). However, to date, it has not been feasible to implement more than a temporary halt in the progression of the disease in specific cases. Because the agents only showed their potential benefits in people with a prodromal or mild stage of Alzheimer's disease. The agents therefore appear to be a step in the right direction rather than a game changer. Nevertheless, the efficacy of these pharmaceutical agents exemplifies the importance of two cornerstones currently employed in the fight against dementia: **Delay** and **early detection**. While a cure for the disease still seems to be out of reach, the progression of the disease can be delayed, and the cognitive decline and associated symptoms can be reduced. A promising delay in the progression of Alzheimer's disease can only be achieved by intervening at earlier stages of the disease, which highlights the importance of early detection.

The objective of this thesis is to identify (new and modifiable) risk factors and explore their relations to the trajectories of the risk of dementia. The results contribute to supporting the prevention of

dementia and facilitating the early detection of such cases, thereby delaying the progression of the disease. Prevention should help to extend the number of healthy life years and to minimize the duration of periods of disability. For effective prevention, it is essential to view the identification and reduction of risk factors as a continuous challenge throughout the entire life course. The overarching question of this thesis was addressed in four studies, each of which posed a specific research question.

Study 1

The potential of routinely collected data to predict dementia risk has not been fully exploited. However, data types such as health claims data are easily and cost-effectively available. Data-driven approaches, such as machine learning methods, have their strengths in high-dimensional data (Breiman 2001b) and may therefore be advantageous for predicting dementia risk from claims data. Individual dementia risk predictions from claims data using data-driven approaches can contribute to the early detection of dementia and the identification of established and potentially unknown risk or protective factors associated with dementia.

Previous reviews have highlighted the need for further models in different populations (Hou et al. 2019; Tang et al. 2015; Goerdten et al. 2019). In recent years there has been an increased focus on routinely collected health data to address this limitation (Ben Miled et al. 2020; Ford et al. 2018; Jain et al. 2020; Nori et al. 2019; Albrecht et al. 2018). Although the German healthcare system is one of the largest in the world, German claims data has hardly been used to predict dementia.

Accordingly, the objective of this study was to investigate the following three questions: **Are German claims data suitable for individual dementia risk predictions? Can machine learning outperform classical regression methods in terms of predicting dementia risk? Which features are important predictors of dementia, and can new features be identified in addition to established risk factors?**

Study 2

The motivation for Study 2 was to gain a deeper insight into the pathways between education, diabetes and cognitive function. The exploration of these associations may contribute to the identification of groups that are particularly vulnerable to cognitive decline and dementia.

Previous research has examined the combined effects of education and diabetes on cognitive performance. The findings indicated that individuals with diabetes exhibited poorer cognitive performance, and even more so if they had lower levels of education. Kowall & Rathmann concluded that the effects were additive (Kowall and Rathmann 2023). A further study examined the pathway between socioeconomic status and dementia by evaluating lifestyle-related diseases as potential mediators. However, the results did not yield a significant association between education and the risk

of diabetes. Consequently, a conclusion about the role of diabetes as a potential mediator was limited (Nakahori et al. 2018).

Given that low education and diabetes are well-established risk factors for cognitive decline and dementia, and the precise mechanisms by which they are associated with cognitive function are only partially understood, further investigation was needed. Therefore, the question is derived **whether diabetes may act as a potential mediator on the pathway of education and cognitive function.**

Study 3

In order to provide evidence for the development of potential interventions, or in this case, the timing of such interventions, it was necessary to gain a deeper insight into the dementia risk trajectory associated with diabetes.

The Lancet Commission report from 2020 still classified diabetes as a late-life risk factor (Livingston et al. 2020), whereas the latest report from 2024 categorized it as a risk factor for mid-life (Livingston et al. 2024). This illustrates that the association between diabetes and dementia continues to be a topic of great research interest and need to be further discussed. Previous studies reported an elevated risk of cognitive decline in individuals with prevalent diabetes, rather than in those with incident diabetes (Croxson and Jagger 1995) or in those with a disease duration of up to 8 years (Wu et al. 2015). Furthermore, studies that emphasize the role of age at the onset of diabetes have provided evidence that the risk of developing dementia increases with the duration of diabetes (Barbiellini Amidei et al. 2021; Xu et al. 2009).

The underlying mechanisms through which diabetes affects the risk of dementia, particularly in the initial phase following the onset or diagnosis of diabetes, remain a partially inconclusive area of research. Accordingly, this study examined the following: **What is the temporal trajectory of the risk of developing dementia in individuals diagnosed with diabetes?** Further examinations addressed the impact of diabetes severity and treatment strategy on this relationship.

Study 4

The identification of risk factors also requires a critical examination of the nature of the association between exposure (late-life depression) and outcome (dementia). This entails establishing whether a presumed risk factor is, in fact, a risk factor that can be addressed through an intervention approach. A factor that increases the risk of a disease may also be a prodrome, i.e. an early symptom of a disease that has not yet been detected. Although a prodrome does not directly serve as a starting point for prevention, it can nevertheless contribute to the early detection of the disease. Examining of the risk trajectory of incident late-life depression and subsequent dementia may contribute to a more comprehensive understanding of the nature of this association.

On the one hand, previous research concluded that the nature of the mechanism between late-life depression and dementia is a prodromal one (Singh-Manoux et al. 2017; Mirza et al. 2014; Almeida et al. 2017; Tapiainen et al. 2017). On the other hand, there is evidence for the assumption that depression is a risk factor for dementia (Wu et al. 2020; Cherbuin et al. 2015; Byers and Yaffe 2011).

The question of **whether late-life depression is a risk factor or a prodromal symptom of dementia** remains a topic of ongoing scientific debate. This is also demonstrated by the latest report from the Lancet Commission, which classified depression as a risk factor only in mid-life, whereas the previous report from 2020 contextualized depression as a risk factor in late-life. To investigate this question further, the temporal association between incident late-life depression and incident dementia was explored in Study 4.

Summary of the research questions

- Are German claims data suitable for predicting the risk of dementia and for identifying risk factors? (Study 1)
- Can diabetes be identified as a mediator on the pathway of education and cognitive function? (Study 2)
- How does diabetes duration impact the temporal trajectory of the risk of dementia? (Study 3)
- Can late-life depression be defined as a risk factor or a prodromal symptom of dementia? (Study 4)

Outline

This thesis is based on four studies and is structured in seven chapters. Following this introduction, the outcomes and exposures of the underlying studies are introduced and defined in the theoretical background section 2. The potential for modifying the identified risk factors and their possible impact on the prevention of dementia are also addressed there. The subsequent section introduces the life course approach that builds the theoretical framework of this thesis and the studies on which it is based. Chapter four describes the data sources and statistical methods used in the underlying studies. This is followed by a summary of the main findings of the underlying studies in chapter five. The discussion summarizes the findings from the underlying studies of this thesis, contextualizing them within the overarching objective, the theoretical framework and with open points that need to be discussed. Furthermore, this chapter presents a discussion of the strengths and limitations of the studies, as well as offering perspectives on future research. The thesis closes with a conclusion.

2. Theoretical background

2.1. Dementia

Dementia is an umbrella term for a progressive decline across various domains, including cognitive abilities, psychological well-being, behavioral patterns, and even physical functioning. Dementia is not a singular pathological condition or a single disease. Rather, it is a syndrome with aggregation of characteristics that differs from the characteristics observed in healthy brain aging (Nelson et al. 2011). The consequences of dementia are a gradual deterioration in the ability to perform essential daily activities that are necessary for independent living (Sachdev et al. 2014). The course of the disease is usually progressive with dynamic symptoms (DGN & DGPPN 2023).

Dementia Subtypes

As already mentioned, dementia is an umbrella term for a range of disease with group of common symptoms, also known as dementia subtypes. The most common dementia subtype is Alzheimer's disease (Cao et al. 2020; Goodman et al. 2017). The principal characteristic of Alzheimer's disease is the accumulation of beta-amyloid and tau proteins in the brain, which results in the impairment of cognitive functions. The typical resulting symptoms are alterations in memory, abstract thinking, judgment, behavior, mood, and emotions, as well as the disruption of physical coordination (Gauthier et al. 2021). The second most common dementia subtype, Vascular dementia (Cao et al. 2020; Goodman et al. 2017), often manifests as a consequence of the brain being unable to receive the essential nutrients and oxygen supplied by cerebral blood circulation. This phenomenon can occur subsequent to a single stroke in a vital cerebral region, or as a result of multiple minor cerebrovascular events (Gauthier et al. 2021). Further subtypes with a considerably lower prevalence are dementia with Lewy bodies and frontotemporal dementia. The former is characterized by symptoms similar to those observed in Alzheimer's disease, with the additional presence of diminished motor skills, which can make the diagnosis challenging. The latter subtype is accompanied by atrophy of the frontal and temporal cortexes and entails alterations in speech, personality, behavior, impulse control, and coordination (Gauthier et al. 2021). There exist other subtypes of dementia that occur in connection with other diseases (e.g. dementia in Parkinson's disease)(Hanagasi et al. 2017). There are also so-called mixed dementias, in which different dementia subtypes coexist (Fierini 2020).

Global Epidemiology

The number of people living with dementia are estimated up to 57.4 million people with a female-to-male-ratio of 1.67 (Nichols et al. 2022). The prevalence may be notably underestimated in low- and middle-income countries (Cleret de Langavant et al. 2020). Europe and North America showed a higher

prevalence rate than Asia, Africa, and South America (Cao et al. 2020). For an overview of regional differences see (World Health Organization 2021).

Epidemiology in Germany

The dementia prevalence in Germany in 2014 was estimated to 10.3% of the individuals aged 65+, based on health claims data. This is analogous to 1.8 million individuals. By the age of 80–84 years, 13–16% of individuals are affected, and 49–58% of those aged 100+. Similarly, the incidence increased, from 0.4 newly diagnosed cases per 100 person-years in the 65–69 age cohort to approximately 11.9 in individuals aged 95 and above. There is a greater prevalence of the condition among females than males, which may be attributed, at least in part, to the fact that the life expectancy of women is higher than that of men. The most prevalent subtype is Alzheimer's disease (60–80%), followed by vascular dementia (5–10%), with mixed dementia being more prevalent in the geriatric population (Georges et al. 2023; DGN & DGPPN 2023).

Estimations for Germany based on data from the WHO Global Status Report and the Federal Statistical Office showed similar results for 2021. The prevalence was estimated to 1.7 million individuals aged 65+, and 360,000 to 440,000 new dementia cases per year, in the corresponding age group (Blotenberg et al. 2023).

Although there are indications of a decline in the incidence of dementia in Germany (Doblhammer et al. 2015; Wolters et al. 2020), the ageing of society, particularly the fact that the baby boomers will reach the age of 65+, is likely to result in an increase in the number of people with dementia in Germany. (Georges et al. 2023; DGN & DGPPN 2023; Blotenberg et al. 2023).

Pathogenesis

From a pathogenesis perspective, all types of dementia are characterized by a number of common features and neurodegenerating processes in the brain (Elahi and Miller 2017). Furthermore, a number of molecular brain mechanisms are shared among all forms of dementia, including oxygen deficiency (hypoxia), oxidative stress, neuroinflammation, neurodegeneration and alterations in blood–brain barrier permeability. The neuropathological characteristics associated with the different subtypes of dementia are beyond the scope of this thesis and are therefore outlined in the following papers: (Raz et al. 2016; Arvanitakis et al. 2019).

Diagnosis

The dementia syndrome is diagnosed on the basis of the clinical examination, including the psychopathological findings, the medical history and using cognitive tests (DGN & DGPPN 2023; Arvanitakis et al. 2019). The following cognitive or behavioral neuropsychiatric symptoms represents

clinical core criteria for all-cause dementia: Impairment in the ability to function at work or in usual activities; A decline from previous levels of functioning and performing that are not explained by other disorders; The detection of cognitive impairment through patient's medical history and a mental status examination or neuropsychological testing; The cognitive or behavioral impairment in two of the following domains: memory, reasoning and judgement, visuospatial abilities, language functions, change in personality and/or behavior (McKhann et al. 2011). The most common test is the Mini-Mental-Status-Examination (MMSE) (Arevalo-Rodriguez et al. 2015). Biomarkers and imaging techniques are employed for the diagnosis of subtypes and for differential diagnosis (DGN & DGPPN 2023). Clinical assessments, differential diagnostics, and oversight of dementia predominantly transpire in a primary care setting, with suitable specialist contributions as required (Arvanitakis et al. 2019).

With reference to the latest version of the International Classification of Diseases (ICD-11) dementia is defined by significant deficits in multiple cognitive domains that deviate from expected cognitive performance based on age and premorbid functioning, indicating a deterioration from the individual's prior capabilities. Although memory impairment is the most commonly observed deficit, cognitive deficits extend beyond memory and encompass executive function, attention, language, and other domains with potential neurobehavioral changes that may initially manifest as symptoms (World Health Organization 2022).

Definition of dementia in the Studies 1, 3 & 4

In this thesis as well as the Studies 1, 3 & 4, the definition of dementia was based on the ICD-10 system (World Health Organization 2004) because this was the current version of the classification system when the corresponding data has been drawn. A clear differentiation between the subtypes is often challenging to determine. This is presumably one of the reasons why high proportion (45%) of dementia diagnoses (in Germany) are coded as unspecific dementia (ICD-10: F03) (Fink 2014). Accordingly, in this thesis and the included studies a definition of all-cause dementia was employed. The following ICD-10 codes were used to define dementia: dementia in AD (F00), without dementia in AD with early onset (F00.0); vascular dementia (F01); dementia in other diseases classified elsewhere (F02); unspecified dementia (F03); delirium superimposed on dementia (F05.1); and AD (G30), without AD with early onset (G30.0).

2.2. Cognitive function (decline)

Cognitive function, cognition or cognitive abilities are terms that denote a multifaceted concept that encompasses the mechanisms involved in the processing of information in the brain. The concept

includes the domains of perception, memory, learning, attention, decision making, executive control, and language abilities. Metaphorically, the cognitive function of humans can be conceptualized as the software of a computer where the brain is analogous to the hardware (Kiely 2014).

The impairment of cognitive function is a main symptom of dementia. Analyzing cognitive function (or cognitive impairment) as an outcome with respect to protective or risk factors can contribute to the prevention and early detection of dementia before it becomes a clinical manifest stadium. The clinical manifestation of a cognitive impairment that does not (yet) fulfil the criteria for a dementia diagnosis is called mild cognitive impairment (MCI). It is a preclinical and transitory stage between healthy ageing and dementia (Bai et al. 2022). However it is preclinical, MCI can be diagnosed or coded according to the criteria of the ICD classification (World Health Organization 2022).

Definition of cognitive function in Study 2

In Study 2 cognitive function was measured by tests from the Cogstate Brief Battery, a computer-administered cognitive test battery that measure psychomotor function, attention, working memory and memory. The test battery has previously been used in other studies to detect mild cognitive impairment and cognitive impairment in Alzheimer's disease (Maruff et al. 2013; Lim et al. 2013). The Cogstate Brief Battery includes tests to the cognitive domains of psychomotor function, attention, visual learning, and working memory.

2.3. Risk factors

A risk factor is characterized by an increased likelihood of occurrence of a subsequent disease or other health-related outcome. Risk factors are characteristics of individuals that can be used to divide a population into high risk and low risk subgroups. From a public health perspective, risk factors are of particular relevance in the identification of individuals for primary preventive measures and in the evaluation of the efficacy of prevention programs in controlling the targeted risk factors. The influence of risk factors on disease outcomes may be direct or indirect (Kirch 2008; Offord and Kraemer 2000). Risk factors can be categorized into modifiable factors that are potentially beneficial targets for prevention of a disease (e.g.: education, diabetes, physical inactivity) and non-modifiable risk factors (e.g.: sex, age, genetics) (Offord and Kraemer 2000). Protective factors are the antithesis of risk factors, in that they are characterized by a decreased likelihood of occurrence of a subsequent disease or other health-related outcome.

The categorization of a factor as a risk factor is influenced by a number of factors, including age, specific time periods and contextual factors. This indicates that potential risk factors should be considered

within the context of the life course. For example, the occurrence of diabetes before the age of 70 is clearly associated with an increased risk of dementia, whereas the same cannot be determined for cases occurring after this age. (Barbiellini Amidei et al. 2021).

2.3.1. Education: A cohort-related modifiable risk factor?

The connection between educational attainment and cognition has been the subject of considerable research (Panico et al. 2023) and is supported by the theoretical framework of cognitive reserve. The accumulation of educational qualifications and the experience gained from occupational complexity over the life course serve to reinforce resilience against the neurodegenerative processes associated with age-related brain changes, and delay the onset of cognitive decline or the clinical stage of dementia (Stern 2012, 2009). Even if education is only one component in the accumulation of cognitive reserve over the life course, educational attainment has been identified as the most reliable proxy indicator for cognitive reserve (Peña-González et al. 2020). Although the cognitive reserve is an established and well-evidenced theory for the relationship between education and cognition (Nelson et al. 2021; Pettigrew and Soldan 2019), the exact mechanisms by which education affects cognition in later life are complex and remain partially unclear (Herd and Sicinski 2022; Lövdén et al. 2020; Seblova et al. 2020). Moreover, empirical evidence indicates that educational attainment may also interact with the levels of tau protein accumulation within the brain and its implications for cognitive function and dementia (Yasuno et al. 2020).

In the context of dementia prevention, a key question is whether the risk factor of education can be modified and, if so, what potential this may offer for developing strategies to fight dementia. Once an individual has reached middle age or old age, it is hardly possible any more to modify this risk factor (from a public health perspective). Even if an increased educational attainment later in life may also be protective (Jin et al. 2024), it is unlikely that the positive effect in terms of cognitive reserve that accumulates over the life course can be attained to the same amount (Soh et al. 2024; Hayes-Larson et al. 2023). Nevertheless, the situation differs when education is considered a cohort-related modifiable risk factor. In such a case, the issue of education would be solely addressed at an early stage of life. Evidence from the US and China suggested that varying degrees of changes in educational achievement partly explain the disparities in dementia incidence and prevalence between sub-populations (Liu et al. 2022; Hayward et al. 2021; Ding et al. 2020). Evidence from France suggests that an increased educational attainment may be even associated with a reduction in the incidence of dementia (Pérès et al. 2017; Grasset et al. 2016). Gao and colleges find similar results for the role of education with respect to the incidence of cognitive impairment in China (Gao et al. 2017).

Definition of education in Study 2

The definition of education in Study 2 was based on individual's level of educational attainment according to the highest obtained degree, as self-reported by the participants. Higher education was categorized as either higher vocational education or university education. (Reinke 2024).

2.3.2. Diabetes: An individually modifiable risk factor?

In this thesis and the associated original studies, the term "diabetes" is used to refer exclusively to type 2 diabetes mellitus (ICD-10: E11).

Type 2 diabetes mellitus is a chronic metabolic disorder characterized by persistent hyperglycemia, a pathological state defined by an elevated level of glucose in the blood. The underlying pathophysiological defects in type 2 diabetes are insulin resistance in muscle and liver, as well as β-cell failure which is accompanied by impaired insulin secretion. The etiology of the disease is multifactorial including genetic, lifestyle (diet, physical inactivity and obesity), socioeconomic and environmental factors, and aging. Diabetes is associated with microvascular complications (retinopathy and neuropathy) and macrovascular complications, leading to an increased risk of cardiovascular diseases (Bellou et al. 2018; DeFronzo et al. 2015; DeFronzo 2009). Vascular complications and the alterations in glucose and insulin as well as amyloid metabolism are the pathophysiological mechanisms that makes diabetes to a risk factor for dementia. However, precise mechanisms underlying this associations remain unclear (Xue et al. 2019; Biessels et al. 2006; Ott et al. 1999). The current understanding of the association is multifactorial in nature and involving a range of factors including cardiovascular risk factors, glucose toxicity, changes in insulin metabolism, as well as inflammation (Ninomiya 2019). Furthermore, there is an existing correlation between diabetes and a reduction in brain volume, as well as an increase in brain atrophy over time (Zhang et al. 2022).

Similar to dementia, diabetes is a chronic disease with progressive characteristics. The progressive nature of the disease implies a changing risk for the development of complications, which is dependent on both the age at onset and the duration of diabetes (Zoungas et al. 2014). The same can be applied to the association with the risk of dementia (Li et al. 2021), at least if the onset is before the age of 70 (Barbiellini Amidei et al. 2021). Similarly, there is usually a considerable time interval between the initial onset of symptoms and the subsequent diagnosis (Harris et al. 1992). Nevertheless, diabetes is an individually modifiable risk factor for dementia in two respects. On the one hand, the risk of developing diabetes can be reduced through early detection (prediabetes) and appropriate prevention in the form of lifestyle interventions. The main factors here are diet and physical activity (Walker et al. 2010; Uusitupa et al. 2019). On the other hand, glucose management and the associated blood glucose

settings in people with diabetes can also be seen as a modifiable risk factor. This is due to the fact that effective glycemic control mitigates the risk of diabetes-related complications (e.g. retinopathy, neuropathy, peripheral artery disease), which represent a crucial pathway for the association between diabetes and dementia (Zheng et al. 2021; Chang et al. 2021). The earlier diabetes is diagnosed, the earlier appropriate measures can be taken to achieve optimal glycemic control and thus reduce the risk of complications and subsequent diseases, including dementia.

In 2021, statistical projections indicated that the global prevalence of diabetes has reached approximately 537 million individuals. This number was expected to reach an estimated 783 million people living with diabetes by 2045 (Sun et al. 2022). This highlights the relevance of diabetes as a modifiable risk factor for cognition and dementia. This was further emphasized by a study which demonstrated that 9.9% of dementia cases in a population aged 70-74 were attributable to diabetes (Rolandi et al. 2020).

Definition of diabetes in Study 2 & 3

In Study 2, diabetes was defined by the occurrence of a minimum of one of the following conditions: Self-reported diagnosis of diabetes, HbA1c ≥ 6.5%, fasting plasma glucose ≥ 7 mmol/L, random plasma glucose ≥ 11.1 mmol/L or use of any medical diabetes treatment. In Study 3 diabetes was identified by ICD-10 code E11 with a validation procedure that is described in section 4.2.3.

2.3.3. Late-life depression: Risk factor or prodromal symptom?

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), a major depressive disorder can be characterized by the presence of depressed mood, markedly diminished interest or pleasure in all, significant changes in weight or appetite, insomnia or hypersomnia, psychomotor agitation, fatigue or loss of energy, feelings of worthlessness, diminished ability to think or concentrate, and recurrent thoughts of death (American Psychiatric Association 2013). A definition of late-life depression is given by a major depressive episode occurring usually after the age of 60 or 65 years. Late-life depression may manifest as either a late-onset condition or as a recurrence or continuation of an earlier-onset depression in later life. It is a psychological disorder that is frequently associated with cognitive impairment and dementia (Butters et al. 2008). Nevertheless, the relationship between late-life depression and cognitive function and dementia remains complex and incompletely understood, with evidence suggesting a bi-directional association (Husain-Krautter and Ellison 2021; Hsiao and Teng 2013; Reynolds et al. 2019). Late-life depression and dementia share a few etiological factor like inflammatory processes, cardiovascular and cerebrovascular dysfunctions (Husain-Krautter and Ellison 2021; Byers and Yaffe 2011). There is evidence for a pre-disease state that

is shared by depression (not only late-life depression) and vascular cognitive impairment, prior to the clinical manifestation of both (Sapsford et al. 2022). The bi-directional association between depression and dementia is still an open question with evidence for depression as a risk factor for dementia (Wu et al. 2020; Cherbuin et al. 2015; Diniz et al. 2013) as well as a prodromal symptom of dementia (Almeida et al. 2017; Tapiainen et al. 2017; Mirza et al. 2014). The association is supposed to be time-dependent in a way that the timing of depression onset is crucial for the subsequent risk of dementia. On the one hand, depression in early and mid-life appears as an important risk factors for dementia. On the other hand, the association of late-life depression with dementia is still unclear (Marawi et al. 2023; Sinclair et al. 2023; Elser et al. 2023). Moreover, the lancet commission classified depression as a late-life risk factor in 2020 (Livingston et al. 2020) and contextualized it to the mid-life risk factors in 2024 (Livingston et al. 2024).

Definition of late-life depression in Study 4

The following ICD-10 codes were used for the identification depression diagnoses: depressive episode (F32); recurrent depressive disorder (F33); and dysthymia (F34.1). But, F33 and F34.1 were used only for validation of F32 diagnoses (Heser et al. 2020). The validation procedure is described in section 4.2.4. Given that the study population consisted of individuals aged 65 and older, late-life depression is defined in this context as late-onset depression.

3. Theoretical framework: Risk factors for dementia over the life course

The objective of this chapter is to provide a common theoretical framework for the thesis, integrating the individual research questions with the overarching topic of the thesis. This chapter focuses on Studies 2-4, as Study 1 was based on a data-driven approach with an explorative character. Although Study 1 is based on a different study design the life course approach presented here may also be useful for the subject of the first study, i.e. predicting the risk of dementia by considering information from all phases of life.

Dementia is generally attributed to the late-life. Actually, the dementia prevalence in people younger thang 65 (young-onset dementia) is low (around 119 per 100,000 people) (Hendriks et al. 2021). However, the risk factors for dementia, particularly those that can be modified, offer opportunities for intervention and prevention over the entire life course.

According to the report by the Lancet Commission, 14 modifiable risk factors, are distributed across the entire life course, encompassing the periods from early to mid- to late life (see Figure 2). These include low education, diabetes as well as depression (Livingston et al. 2024). This report highlights that the challenges and intervention points for fighting dementia (from a public health perspective) can be found throughout the entire life course. Furthermore, individual opportunities for prevention should be seen as lifelong prospects for reducing the risk of dementia. Therefore, the concept of life course epidemiology also known as the life course approach provides a suitable framework for embedding the objectives and scope of this thesis.

Life course epidemiology is a framework comprising elements from a range of scientific domains, including epidemiology, sociology, psychology, biomedical sciences, and other fields related to population health sciences (Wagner et al. 2024). The framework considers the timing and duration effects of social and physical exposures and risk factors as well as their accumulation and interaction for health outcomes over the entire life. The life course encompasses the life stages during gestation, childhood, adolescence, young adulthood, and later adult life. The purpose is to model pathways that link exposures over the life course to health outcomes in later life. By including age at onset or trajectory, transition and turning points of health behavior and diseases, it also opens a temporal perspective on risk factors that appear or manifest in adulthood or later (Kuh et al. 2003; Ben-Shlomo and Kuh 2002; Lynch et al. 1997). The framework of life course epidemiology includes the following concepts: “**Accumulation of Risk**”, “**Chain of Risk Model**”, “**Mediating Factors**”, “**Trajectory, Transition and Turning Points**”, “**Time**” and “**Induction and Latency Periods**”.

The objective of Study 2 was to ascertain whether the relationship between education and cognitive function is partially linked via diabetes. Three of the above-mentioned concepts can be used to embed

this question. The hypothesized association of education via diabetes on cognitive function can be described as a chain of risks: Lower levels of education, which are usually formed in early life, lead to a higher risk for diabetes, and diabetes, which manifests in the majority of cases in mid- or late-life, increases the risk for a declined (impaired) cognition in late-life. In this consideration of the hypothesis as a pathway model, as well as in the consideration of the question itself, diabetes appears as a mediating factor between the exposure and the outcome, whereby the sequential pathways that link the individual factors with each other cross the life stages. However, both factors may also be independently associated with the outcome, independent of the sequential pathways. From another point of view, the potential effects of education and diabetes on cognitive function may be interpreted in an additive way. Thus, lower education and diabetes are different exposures that accumulate risk of cognitive decline over the life course.

Study 3 questioned the temporal evolution of the risk of dementia in individuals diagnosed with diabetes. The hypothesized duration dependency of the dementia risk represents a special case of risk accumulation. In addition, the duration of a disease is a specific concept of time, that usually runs equivalent to individual's lifetime. In contrast to lifetime, however, disease duration must be clearly separated from age, as duration effects can differ at different ages. A further perspective is that the trajectory of dementia risk after a diabetes diagnosis can be seen as a transition from one risk pattern to another one (e.g. a decreasing to an increasing risk pattern), whereby this change can be interpreted as a turning point.

The underlying question in Study 4 was if late-life depression is a risk factor or a prodromal symptom for dementia? This question was explored by analyzing the trajectory of dementia risk, but in subsequence of a depression diagnosis. However, it was not the duration of the disease that mattered, but the time since diagnosis. This is a different concept of time because depression is not necessarily a chronic condition, but diabetes is. If late-life depression is actually an early symptom of dementia that occurs before the clinical status manifests, it represents a transition from a pre-clinical to a clinical state. Evidence for a prodromal association may be indicated by a greater strength of association for shorter time intervals between incident late-life depression and subsequent dementia. This question, together with the circumstance that in the case of dementia can often be long phases between the onset of the disease (the process) and the diagnosis (Karr et al. 2018), means that this is a latency period. Late-life depression as an early symptom therefore falls into a latency period for dementia that has not yet reached a clinical stage.

One if not the elementary component of the life course approach is time, which is not limited to life-time. For this reason, longitudinal studies are fundamentally necessary for studies related to the life course approach. Preferably, although not necessarily, these should be of a prospective design.

The lack of a standardized and strict definition of life stages is a consequence of the fact that such a classification should always be based on the specific context of the research question. A orientation rather than a definition of the periods of life based on Levinsons Model of Development in Early and Middle Adulthood (Cited from: Dannefer 1984). **Early-life** encompasses the childhood and adolescence phase between 0-17 years as well as the early adulthood between 17-40 years; **Mid-life** includes the phase of middle adulthood from 40-60 years; The **Late-life** is defined as the late adult transition from 60-65 years, and the late adulthood denotes the ages from 65 and upwards. This categorization is also appropriate for the purposes of this thesis, in relation to the underlying exposures and outcomes. The process of educational acquisition usually takes place in early life. The global incidence of diabetes shows a peak among the age group 55-59 (Khan et al. 2020), indicating that this risk factor is mainly located in mid-life. However, the mean age at diagnosis in Germany is between 61-63 years (Jacobs et al. 2020), what rather suggests that it is perhaps the transition phase from middle to late-life. The majority of diagnoses of dementia occur in late-life (Wolters et al. 2020; Blotenberg et al. 2024). Late-life depression is per definition located in late-life.

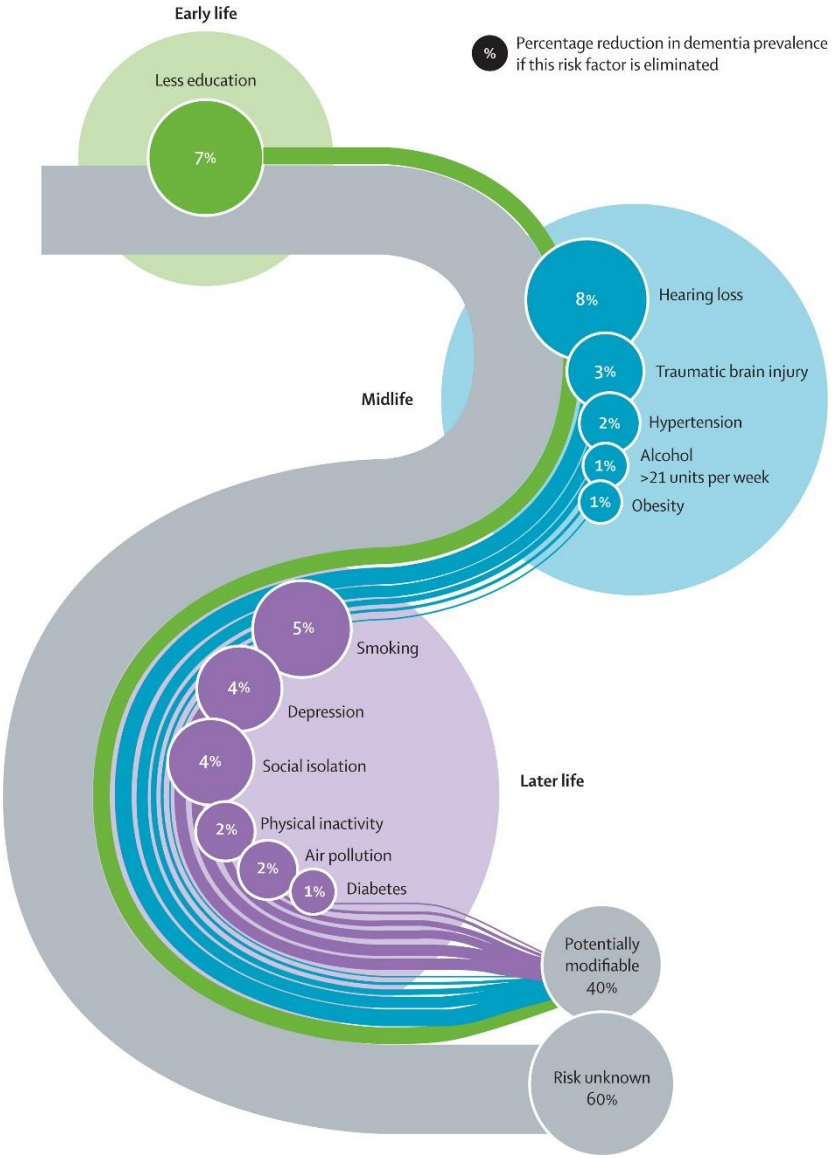
Empirical evidence according to the theoretical framework of life course

Although the foundations of educational attainment typically start early in life, there is an association with the dementia risk in the late life phase (Seblova et al. 2020). Additionally, higher education increases the likelihood for higher and more complex occupations (Hyun et al. 2022) and this may associate with better conditions for the accumulation of a higher cognitive reserve over the life course (Dekhtyar et al. 2015). Thereby, the risk of dementia can be decreased or rather delayed (as described in 2.3.1). Nevertheless, studies have indicated that the underlying mechanistic pathways of education and cognition remain unclear (Seblova et al. 2021). However, it is noteworthy that the risk of dementia is also reduced by cognitively stimulating jobs, independent of the education level (Kivimäki et al. 2021).

Physical inactivity and nutrition are factors that can have a multifaceted influence on the risk of dementia, especially or only in mid-life. Physical activity protects against hypertension, for example, but can also reduce neuroinflammatory processes (Huuha et al. 2022; Freitas et al. 2020). Lifestyle-related diseases like obesity (Tang et al. 2021; Qu et al. 2020), hypertension (Mahinrad et al. 2023; Hughes et al. 2020), high LDL cholesterol (Wee et al. 2023; Zhang et al. 2020) and (Xue et al. 2019; Barbiellini Amidei et al. 2021) suggest a potentially indirect association of nutrition to the dementia risk over the whole life (e.g. control of body weight (Veronese et al. 2017)) (Kjeldsen et al. 2022; Liu et al. 2020). The link of obesity and dementia may operate through diabetes and hypertension as well as the association with cardiovascular diseases (Safaei et al. 2021). The mechanisms linking diabetes and dementia were described in detail in section 2.3.2. Other lifestyle factors such as smoking (Raggi et al.

2022; Gottesman et al. 2017) and excessive alcohol consumption (Kivimäki et al. 2020) in mid-life also increase the risk of dementia. While alcohol induce alterations in the structure and function of the brain (Rehm et al. 2019) smoking and dementia are associated by cardiovascular pathways and through oxidative damage (Choi et al. 2018; Peters et al. 2008). Further important risk factors in midlife are hearing loss (Yu et al. 2024; Wei et al. 2017) and traumatic brain injuries (Gardner et al. 2023; Gu et al. 2022). Traumatic brain injuries can cause dementia in a direct way through a traumata (Katz et al. 2021) or by an increased deposition of amyloid-beta (Aβ) and tau aggregates in the brains (Brett et al. 2022; Graham and Sharp 2019). The hypothesized mechanisms for the connection of hearing loss and dementia incorporate a reduced cognitive reserve (Griffiths et al. 2020) as well as psychosocial factors, such as loneliness, social isolation (Shukla et al. 2020) and depression (Powell et al. 2022). Depression is also independently associated with the risk of dementia (Stafford et al. 2022; Elser et al. 2023). The association was described in detail in section 2.3.3. However, there is only reliable evidence for depression as a risk factor in mid-life (Karlsson et al. 2015). In late-life, being the main focus of Study 4, the evidence is not clear and depression rather appears as a prodromal symptom of dementia (Mirza et al. 2014; Heser et al. 2020).

Figure 1: Population attributable fraction of potentially modifiable risk factors for dementia

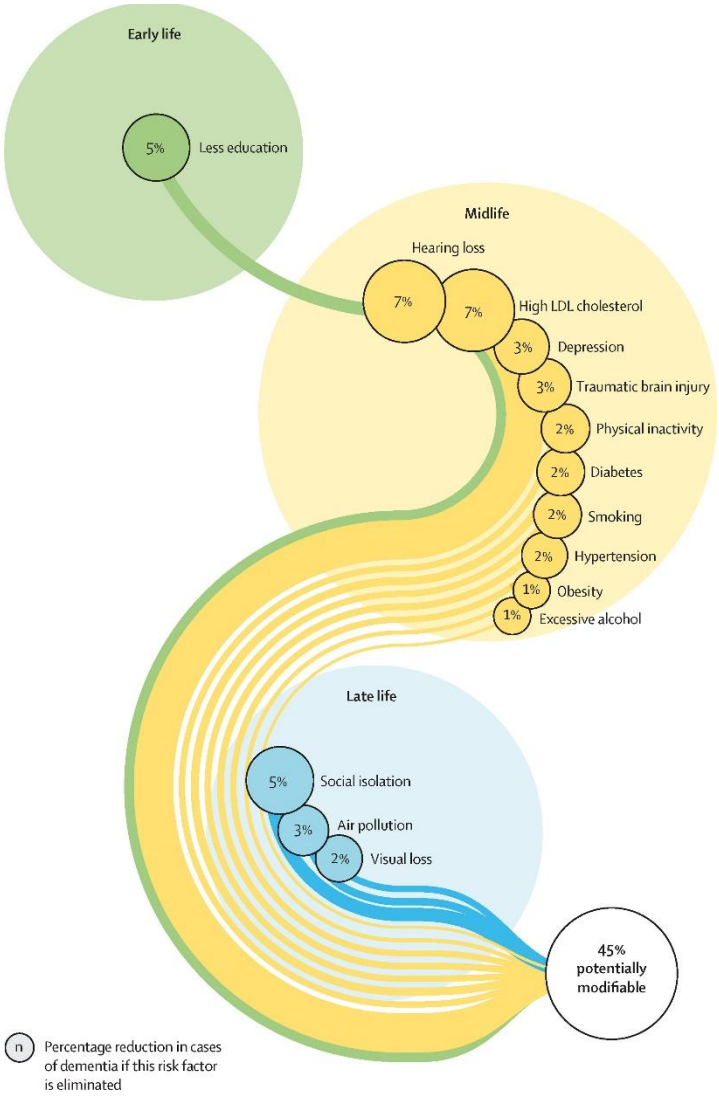


Source: (Livingston et al. 2020)

In the Lancet Commission report from 2020 (see Figure 1), depression was still characterized as a late-life risk factor (Livingston et al. 2020). This reinforces the ambiguity and relevance of this connection.

However, as the classification of the mid-life and late-life stages is not a matter of clear-cut periods and depression can be a chronic disease, a definitive allocation to one of the life stages is challenging. Nevertheless, the definition of late-life depression, namely depression with an onset in late life, provides a clear classification into this phase of life. The status is more clearly defined for the factors social isolation (Elovainio et al. 2022; Shen et al. 2022), air pollution (Peters et al. 2021; Abolhasani et al. 2023) and visual loss (Shang et al. 2021; Kuźma et al. 2021). The mechanism between social isolation and dementia remains largly unclear (Evans et al. 2019; Kelly et al. 2017). Air pollution like fine particulate matter seems to be linked to dementia via lung function and inflammatory processes (Aretz et al. 2021; Jayaraj et al. 2017). There is evidence to suggest that visual loss may be linked to dementia via underlying comorbidities, such as diabetes, as well as via a direct path (Shang et al. 2023; Paik et al. 2020).

Figure 2: Population attributable fraction of potentially modifiable risk factors for dementia



Source: (Livingston et al. 2024)

Life course epidemiology represents a fundamental concept for understanding the evolution of risk profiles for chronic diseases like dementia, thus providing a basis for the development of effective public health strategies (Wagner et al. 2024).

4. Data & Methods

The objective of this chapter is to outline the materials and methods employed in the underlying studies of this thesis. The following section outlines the data sources utilized in the studies. Subsequently, the designs and statistical methods employed in the individual studies are illustrated.

4.1. Data

The studies presented in this thesis are based on two distinct data sources. In Studies 1, 3 and 4, the database was provided by German health claims data from the health insurance company "Allgemeine Ortskrankenkasse (AOK)", which serves as an illustrative example of routinely collected data. The data source for Study 2 was the Lifelines cohort study, a longitudinal population-based health survey from the Netherlands. In the following two sections, the data sources as well as the general pros and cons of the routinely collected data and population-based health surveys are described.

4.1.1. AOK-Data

Health claims data is administrative data that is collected by health insurers on a routine basis for the purpose of maintaining their billing systems and for the purpose of capturing any kind of healthcare utilizations or encounter with the healthcare system (Cadarette and Wong 2015). In Germany, there is a statutory obligation to be insured with a health insurance fund. Therefore, the entire German population is insured by the public (statutory) or private health insurance system (Döring and Paul 2010), whereby approximately 88% are in the statutory health insurance system (GKV-Spitzenverband 2024). Health claims data comprises a wide range of detailed clinical information, including diagnoses, prescriptions filled, and conducted operations and procedures on an individual level. This information is drawn from both outpatient and inpatient sectors. The high dimensionality of information and the possibility of conducting analyses on large sample sizes over long observation periods (with short intervals) make health claims data a valuable resource for research in a number of fields, including epidemiology, health services research, drug utilization and public health. German health claims data has been utilized in numerous studies within these disciplines (Kreis et al. 2016; Hoffmann 2009; Hennessy 2006; Schneeweiss and Avorn 2005), and has also been for prevalence and incidence estimations (Georges et al. 2023; Nerius et al. 2020; Doblhammer and Barth 2018; Nerius et al. 2017; Doblhammer et al. 2015).

The routine collection of health claims data eliminates the need for expensive and time-consuming data collection processes. Consequently, they constitute the foundation for cost-effective

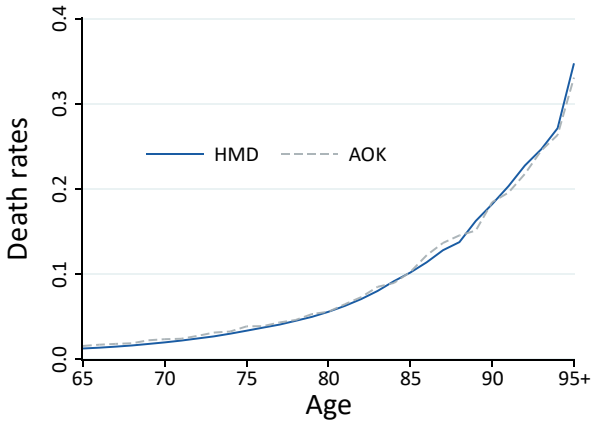
epidemiological and public health-related analyses. Given the administrative nature of the data, participation or inclusion in the data is not optional, and the issue of selection bias is not applicable. Moreover, the German healthcare system provides coverage for all citizens, including those residing in nursing homes, which are often unrepresented in population-based surveys (Gaertner et al. 2019). As the included medical diagnoses are made by healthcare professionals, recall bias, a common issue in interview-based surveys (Coughlin 1990), is not a concern. Information such as diagnoses, prescriptions, and surgeries come from the in- and outpatient sector and are typically coded according to international classification systems, allowing for cross-country comparisons. In the case of death, the data set includes the date of death. However, claims data are not without limitations. Information is generated exclusively when individuals have interacted with the healthcare system. In the context of diagnoses, this can result in undetected cases (with subsequent implications for prevalence estimates) and may be interpreted as a form of reporting bias. As the data is originally collected for the purpose of the billing system, it may be affected by changes to the billing system or to legislation. Further, the data contains no or limited socioeconomic, demographic and geographic information.

If the analysis is restricted to the data of a single insurance fund, as is the case with the studies presented in this thesis, there are additional factors that must be considered. The choice of an insurance company is usually an individual decision, with a few exceptions. It is possible for individuals to change insurance companies, which may result in a loss of follow-up. The distribution of socio-demographic and socio-economic characteristics may differ between insurance companies, resulting in deviations when compared to the corresponding population (in this case, the German population) (Epping et al. 2021). These differences can be particularly large between insurance companies from the public and private system (Greß 2007).

Description of the sample in the underlying studies

Study 1, Study 3 and Study 4 used anonymized health claims data from an age-stratified random sample of 250,000 individuals insured by the largest German health insurance provider, Allgemeine Ortskrankenkasse (AOK). The sample, which was drawn and provided by the AOK Research Institute WIdO (Wissenschtliches Institut der Ortskrankenkassen) in 2004, comprised individuals born prior to 1955, with longitudinal follow-up extending through 2015 (shown in Figure 4). The sample was representative of the German population aged 65 years and older in terms of mortality, as illustrated in Figure 3 (Fink 2017; Fink 2014). The dataset encompassed quarterly information across a multitude of variables, including demographic characteristics (sex, date of birth and date of death), medical diagnoses from both the inpatient and outpatient sector (utilizing the ICD-10 classification system), pharmaceutical prescriptions (in accordance with the ATC classification), and surgical interventions (grounded in the German adaptation of the ICPM classification). The data does not contain electronical health records in terms of reports of symptoms, results of physical examinations or diagnostic test results. Given that the data had been rendered anonymously, there was no necessity for ethical review or patient consent.

Figure 3: Death rates (2004-2007) for the AOK-sample and Germany



Source: Own illustration based on: (Fink 2017)

Validation of dementia diagnoses within the sample

In light of the high proportion of unspecified dementia diagnoses (Fink 2014) (see section 2.1) and the primary use of the data for billing purposes, the question of the validity of the available diagnoses is a consequent issue that must be addressed (Schubert et al. 2010). Although the problems of false-negative and false-positives diagnoses can occur in billing data (Taylor et al. 2009), it has been shown that the specificity for dementia diagnoses is high (Chen et al. 2019; Rizzuto et al. 2018). Nevertheless, to counteract the potential problem of false-positive dementia diagnoses, an internal validation strategy was applied.

The strategy consisted of a two-stage procedure of the diagnosis described in section 2.1. In the initial phase, 'verified' outpatient diagnoses and inpatient discharge or secondary diagnoses were corroborated. In the subsequent phase, diagnoses were substantiated through concurrent in- and outpatient evaluations, multiple diagnoses rendered by various physicians within the same or successive quarters, or the co-occurrence of diagnoses throughout the study duration. In the event of

a dementia diagnosis and death occurring within the same quarter, it was considered to be valid because time-related validation was not possible. This strategy was developed and introduced by Doblhammer and colleges (Doblhammer et al. 2015) and based on Schubert et al. (Schubert et al. 2010). It is well established and repeatedly applied (Doblhammer et al. 2015; Fink 2014; Rakuša et al. 2023; Georges et al. 2023).

4.1.2. Lifelines Data

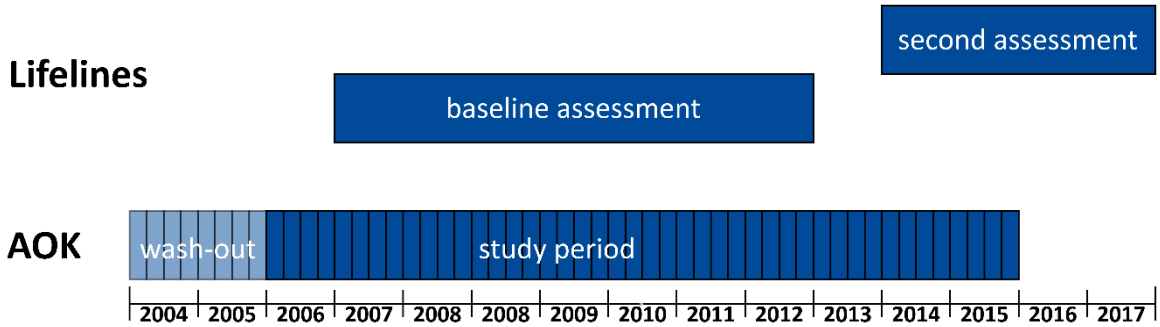
Population-based health surveys are a fundamental instrument in the field of public health and epidemiology research, offering insights into a population's health status, the necessity of health services, and the efficacy of health interventions. They provide valuable data for all fields of medical science and the planning of health programs, they can inform policy makers and serve as a basis for the identification and investigation of risk factors. An example of a population-based health survey is the Lifelines Cohort Study, from which data was used for Study 2.

A principal benefit of population-based health surveys is the comprehensive and extensive range of information, particularly regarding the social, economic and demographic background of the study participants, as well as health-related behaviors or measures of the quality of live. Furthermore, population-based health surveys frequently incorporate medical examinations, allowing for the inclusion of biomarkers, laboratory results, and test scores. This enables the identification of undiagnosed cases and the assessment of disease severity. Additionally, conditions lacking a clinical diagnosis or prodromal phases of certain diseases can be detected. For instance, cognitive tests can quantify cognitive function, which may serve as an indicator for cognitive impairment prior to its clinical manifestation. However, population-based health surveys are not without limitations. In comparison to data that is routinely collected, population-based surveys are both costly and time-consuming to undertake. Typically, population-based surveys do not include individual diagnoses from healthcare professionals; rather, they rely on self-reported information. Such data is typically affected by recall and response bias (Elston 2021). Additionally, population-based surveys are susceptible to selection bias (Tripepi et al. 2010; Strandhagen et al. 2010). This is particularly evident in the exclusion of individuals with severe disabilities as well as the underrepresentation of nursing home residents (OECD 2023). Panel attrition is a further issue that typically occurs in longitudinal population-based surveys (Saiepour et al. 2019).

In the mandatory data statement, Lifelines described the study as follows: “Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North of the Netherlands. It

employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics" (Lifelines 2024).

Figure 4: Overview of the data structure



Source: Own illustration based on: (Sijtsma et al. 2022)

The Lifelines cohort study encompasses a wide range of health-related information, comprising not only results from physical examinations but also detailed analyses of biological samples, comprehensive evaluations from cognitive testing, and a comprehensive questionnaire intended to gather multifaceted data regarding various health-related parameters. The data collection process was extended over a period of seven years, from 2006 to 2013, for the initial baseline assessments. This was followed by a subsequent follow-up phase of data collection between 2014 and 2017, which was dedicated to the second wave of assessments (Scholtens et al. 2015; Sijtsma et al. 2022; Stolk et al. 2008) (shown in Figure 4). Further waves that are part of the ongoing study are the third wave (collection period: 2019-2023) and fourth wave (collection period: started in 2024). These/They are not part of this thesis and the underlying Study 2 respectively.

4.2. Methods & statistical analysis

4.2.1. Study 1: Dementia risk prediction from German claims data using methods of machine learning

The objective of this chapter is to outline the study design and sample as well as the statistical approaches used in the first study (Reinke et al. 2023). The hypothesized questions in this study were if German health claims data suitable for individual dementia risk predictions? Can methods of machine learning outperform classical regression in terms of dementia risk prediction? What are important predictors for the risk of dementia, and is it possible to identify new predictors?

Study design & sample

The research utilized the AOK-dataset, detailed in section 4.1.1. The study period started with 2006 and ended with 2015, with observations recorded on a quarterly basis. The study population comprised individuals aged 65 and older (at baseline) without a dementia diagnosis in either 2004, 2005 or the first quarter of 2006. This resulted in 117,895 individuals, of whom 27,651 received a dementia diagnosis over the course of the study period. In the initial quarter of the diagnosis of dementia, all predictors were assigned the values of the preceding quarter. This was done because the sequence of events within the quarters was unable to be determined. A training population was assembled using a random sample comprising 60% of the study cohort, stratified by dementia status. The residual participants were distributed between a validation and a test population so that the two groups were of equal size.

The definition of the outcome and the respective validation strategy were described in section 2.1 and 4.1.1. The considered predictors were 23 age-related diseases (Doblhammer and Barth 2018; Livingston et al. 2020), all prescriptions coded by the ATC classification system on the level of pharmacological subgroup, and all surgeries based on three-digit classes from Chapter 5 of the OPS, as well as age and sex. With the exception of age and sex, all predictors were incorporated as time-dependent variables. A sum of 324 predictors were considered, with 212 ATC codes and 87 OPS codes.

Statistical methods

In accordance with the data structure (one row per quarter), a discrete process time was defined, starting in the first quarter of 2006. In order to make predictions, logistic regression (LR), gradient boosting machines (GBMs), and random forests (RFs) were employed. This enabled applying discrete hazard models using time-dependent predictors and the accounting for right-censoring (Tutz and Schmid 2016). A LR model was constructed using the combined training and validation population, incorporating all predictors, including categorical process time. This model was regarded as the

benchmark model. A GBM (Zhang and Yu 2005) was trained on the training data with a learning rate of 0.01 and the iterations were limited to 10,000. The stopping criteria was determined by no log-loss improvement over the last 100 iterations. Further parameters were tuned to a grid search on the validation set. The final GBM was fitted on the combined training and validation population. The RF models were trained with 1,000 and 10,000 trees. All other parameters were set to their default values.

The efficacy of the predictive models was evaluated in terms of accuracy, discriminatory power, and calibration. The overall accuracy was quantified using the integrated prediction error, which was modified from the Brier score. This method relies on weighted quadratic differences between predicted and observed survival functions (Tutz and Schmid 2016). The 95% CI were calculated using 5000 bootstrap replications of the test data. To measure the discriminatory power, a time-independent concordance index (C-index) was computed (Schmid et al. 2018; Harrell 1982). Furthermore, calibration plots were analyzed, and an intercept and slope were determined to evaluate potential systematic overestimations or underestimations of predicted risks (van Calster et al. 2019; Miller et al. 1993). To identify the most influential predictive features of the GBM model, a permutation method was employed to derive relative importance scores (Breiman 2001a). For LR, the relevant odds ratios were provided, organized by their absolute z-values.

4.2.2. Study 2: The effect of diabetes in the multifaceted relationship between education and cognitive function

The following section presents a description of the sample, the study design, and the statistical methods employed in Study 2 (Reinke 2024). The underlying research question in this study was if some of the association between education and cognitive function may operate through the diabetes status?

Study design & sample

The Lifelines cohort included a total of 152,860 individuals aged 18 years or older at the initial baseline assessment. Since only 111,959 individuals of the cohort also participated in the second assessment, 40,901 were excluded from the sample. 73,202 individuals who were younger than 50 years at the baseline assessment were also excluded from the sample. Furthermore, 12,220 individuals without the Cogstate examination to measure cognitive function were excluded. This exclusion also applied to those without information regarding educational attainment (120 individuals), the HbA1c-value to define diabetes status (267 individuals), or any relevant confounding variables (19 individuals). As a result, the final sample, which was selected according to the aforementioned inclusion and exclusion criteria, consisted of 26,131 individuals.

In order to implement a longitudinal study design that incorporates a causal time order, the exposure variable for educational attainment (described in section 2.3.1) as well as the mediator diabetes (described in section 2.3.2) and the confounding variables were defined using data from the baseline assessment. As the analysis was conducted only on individuals aged 50 years or older at baseline, it can be assumed that their maximum educational achievement occurred prior to the measurement of the mediators and confounders. The outcome cognitive function was measured only at the second assessment. It was measured by the Cogstate Brief Battery which has been described in section 2.2.

Previous studies have demonstrated that the relationship between diabetes and cognitive function is not consistent across different cognitive domains (Monette et al. 2014; Palta et al. 2014). This evidence indicates that psychomotor functioning and attention are the domains that are most significantly affected by diabetes. Accordingly, a composite score was constructed from these two specific domains, using the results of detection and identification tasks of the Cogstate Brief Battery. This approach was similar to a previous study (Maruff et al. 2013). For this purpose, the log-10-transformed reaction time in milliseconds from both assessments was z-standardized within 5-year age groups, and subsequently aggregated. Consequently, a positive value indicates a comparatively higher (slower) reaction time relative to the corresponding age group, which is indicative of worse cognitive function. Conversely, a negative value indicates a lower (faster) reaction time, which is indicative of better cognitive function relative to the corresponding age group. The confounding variables were the comorbidities: depression, hypertension, stroke, heart failure, and high cholesterol. Further controls were: age, sex, physical activity, obesity, smoking history, monthly income as well as the accuracy of the tasks from the outcome measure defined as the number of correct responses diverted by the number of total responses.

Statistical methods

The relationship between education, diabetes status and cognitive function was investigated using linear and LR models. To assess the indirect effect of the exposure education on the outcome cognitive function via the mediator diabetes, a causal mediation framework was employed. This methodology, which is based on structural equation modelling, enables the estimation of the average causal mediation effect (ACME) also known as indirect effect. The process comprised several steps. Initially, two statistical models were constructed: One for the mediator M (Equation 1) and another for the outcome O (Equation 2). In the subsequent phase, parameters for both the outcome and mediator were simulated from their respective sampling distribution. Thirdly, potential mediator values were simulated prior to assessing the potential outcomes, given the values for the mediator, enabling the calculation of causal mediation effects. Finally, point estimations for the ACME and the direct effect and confidence intervals were derived from the simulated distribution. This methodology is

comprehensively detailed by Imai et colleges (Imai et al. 2010a). The mediator in model 1 was modeled by LR. The outcome was modeled by a linear regression.

$$Model\ 1: \quad M = \beta_0 + \beta_E E + \sum_{i=1}^{n} (\beta_i C_i) + \epsilon_M$$

$$Model\ 2: \quad O = \alpha_0 + \alpha_E E + \alpha_M M + \sum_{i=1}^{n} (\alpha_i C_i) + \epsilon_O$$

A sensitivity analysis was conducted to test the assumption of no unmeasured confounders in the association between the mediator and the outcome (Imai et al. 2010b). This was performed by estimating a correlation parameter for $\epsilon_M \; and \; \epsilon_O$ at which the ACME would be zero (Tingley et al. 2014). Further sensitivity analyses were conducted by: A classical structural equation modeling as an alternative mediation approach, testing the robustness of the study design by using other indicators than diabetes as potential mediators, and an extended definition of diabetes based on information from the baseline and follow-up periods.

4.2.3. Study 3: Diabetes duration and the risk of dementia: a cohort study based on German health claims data

This section presents a description of the sample, an account of the study design, and an exposition of the statistical methods used in Study 3 (Reinke et al. 2022). Study 3 examined the temporal evolution of the risk of dementia in individuals diagnosed with diabetes, as well as the potential modifying effects of diabetes severity and treatment type on this association.

Study design & Sample

The database utilized in this study was the AOK-dataset, which was described in detail in section 4.1.1. The study population comprised individuals with incident diabetes who were 65 years and older at baseline (2006). Individuals with a prior diagnosis of dementia or diabetes, or with any type 1 diabetes diagnosis, were excluded from the study. The cohort consisted of 13,761 individuals aged 65 years and older who had been newly diagnosed with diabetes and had been followed for a minimum of one quarter. Over the course of the observation period, which concluded at the end of 2014, 2,558 cases of dementia were documented, while 2,845 individuals died during the study period. Additionally, 8,544 participants completed the follow-up, and 107 individuals exited the dataset for various other reasons, including changes in insurance company. The outcome of dementia and the corresponding validation approach are described in section 2.1 and 4.1.1.

The diagnoses of diabetes, as outlined in section 2.3.2, were validated by at least two verifiable diabetes diagnoses in distinct quarters. In order to validate the diagnoses of diabetes and dementia also at the end of the study period (end of 2014), data from the year 2015 were also employed. Diabetes treatment were defined into three groups of anti-diabetic medication (ADM) and one group with individuals who received **no-ADM**. The groups were classified according to the Anatomical Therapeutic Chemical (ATC) classification system. Individuals in the insulin group were prescribed medication that corresponded to the ATC code A10A. The group comprising **oral-ADM** was associated with prescriptions classified under the ATC code A10B. Concurrent prescriptions of both insulin and oral-ADM were allocated to the category designated as **mixed-ADM**.

The severity of diabetes was implemented by a proxy measure of diabetes complications. The Adopted Diabetes Complications Severity Index (aDCSI) was developed by Young and colleagues as a tool for predicting the risk of mortality and hospitalization in individuals with diabetes (Young et al. 2008). Subsequently, the tool was adopted and validated for utilization within claims data, irrespective of laboratory results (Chang et al. 2012). Wicke and colleagues undertook an update of the tool, migrating it from the ICD-9 to the ICD-10 version (Wicke et al. 2019). The aDCSI ranged from 0 to 13 and for our analyses it was categorized into five levels (0, 1, 2, 3 and 4+). The aDCSI was implemented as a time-varying variable.

Other covariates in the study were age and sex as well as the comorbidities: hypertension, depression, cerebrovascular diseases, ischemic heart diseases, atrial fibrillation and flutter, obesity and disorders of lipoprotein metabolism.

A strategy proposed by Hernan et al. for the avoidance of biases in treatment assignment was implemented (Hernán et al. 2016; Emilsson et al. 2018).

Statistical methods

To evaluate the risk of dementia in relation to the duration since the initial diagnosis of diabetes, piecewise exponential regression models were employed to explicitly delineate the baseline hazard throughout the study period. The baseline hazard was partitioned into quarters and characterized as a second-degree polynomial function of the time since diabetes diagnosis (d), and a quadratic term (d^2). Second-degree polynomial was chosen because of the best AIC compared to other transformations (logarithmic, exponential, cubic and continuous time without a quadratic term). The hazards of the total duration effect were calculated given the covariate matrix Z by:

$$h(d|\boldsymbol{Z}) = h_0(d)\ exp(\boldsymbol{aZ}), \text{with } h_0(d) = b_1 d + b_2 d^2$$

Treatment strategies were incorporated as a time-varying variable, and to assess the duration effect by treatment groups, interaction effects between the treatment groups (ADM_i) and the baseline hazard function were included. The Index *i* represented the four treatment groups: insulin (*i*=1), oral-ADM (*i*=2), mixed-ADM (*i*=3) and no-ADM which represented the reference group. The hazards were calculated by:

$$h(d, ADM_i|\boldsymbol{Z}) = h_0(d, ADM_i)\ exp(\boldsymbol{aZ}),$$

$$with\ \ h_0(d, ADM_i) = b_1 d + b_2 d^2 + ADM_i(b_{3,i} + b_{4,i} d + b_{5,i} d^2)$$

Sensitivity analyses were conducted to examine distinct age, and severity of diabetes. Moreover, Cox regression models were utilized to statistically confirm that the effects of treatments and diabetes severity on the risk of dementia were independent of the modelling approach.

4.2.4. Study 4: The temporal association between incident late-life depression and incident dementia

The subsequent section presents an overview of the study design, sample, and statistical methods employed in Study 4. The study researched the question if late-life depression is rather a prodromal symptom than a risk factor for dementia. This was examined by testing whether the associations would be stronger for shorter intervals between a late-life depression diagnosis and subsequent dementia.

Study design & Sample

The study employed the AOK-dataset, elaborated upon in section 4.1.1. The study cohort comprised individuals aged 65 and above who had not been diagnosed with depression or dementia in 2004 or 2005. The study cohort included 97,100 individuals, of whom 20,779 received a dementia diagnosis over the course of the study period, which spanned from 2006 to 2014. The definition of late-life depression was described in section 2.3.3. The diagnosis of depression was validated in accordance with the following criteria (Wagner et al. 2018): an inpatient discharge or secondary diagnosis; two outpatient diagnoses in different quarters (within a four-quarter period); or two outpatient diagnoses by separate physicians in the same quarter. Furthermore, if the individual died within the quarter of the initial depression diagnosis. Depression severity is categorized as mild, moderate, severe, or unknown, according to ICD-10 codes. The outcome of dementia and the corresponding validation approach are described in section 2.1 and 4.1.1.

The confounding variables of age, sex, and comorbidities were considered in the analyses. All covariates, with the exception of sex, were considered to be time-varying variables.

Similar to Study 3, a strategy to avoid biases by group assignment was applied (Hernán et al. 2016; Emilsson et al. 2018).

Statistical methods

The incidence rates of dementia and piecewise constant exponential models were calculated. Individuals were censored at the end of the follow-up period in 2014 or at the time of death, whichever occurred first. The analysis time was partitioned in up to 36 quarters in order to facilitate piecewise exponential modelling. The Akaike information criterion was used to assess the suitability of various temporal functions, with the natural logarithm of analysis time identified as the best fir to the data. Cases with the simultaneous diagnoses of depression and dementia were incorporated into the control group.

Table 1: Summary of the study designs of the underlining studies

Study	Research question(s)	Outcome	Exposure (life stage)	Data (age)	Design	Statistical methods
Study 1	Are German claims data suitable for Dementia risk prediction from? Can Machine learning outperform classical regression? What are important predictors for dementia?	dementia diagnosis	no exposure, 324 predictors (all life stages)	AOK-data (65+)	prospective, longitudinal, prediction	logistic regression, gradient boosting machines, random forest,
Study 2	Is diabetes a mediator on the pathway of education and cognition?	cognitive function	education (early life) diabetes (mid-life)	Lifelines (50+)	prospective longitudinal cohort study	linear and logistic regression, mediation analysis, structural equation models
Study 3	Increasing dementia risk with increasing diabetes duration?	dementia diagnosis	diabetes (mid-life)	AOK-data (65+)	prospective longitudinal cohort study	piecewise exponential hazard models, cox regression
Study 4	Is late-life depression a risk factor or prodromal symptom for dementia?	dementia diagnosis	depression (late-life)	AOK-data (65+)	prospective longitudinal cohort study	piecewise exponential hazard models, cox regression

5. Summary of the study results

5.1. Study 1: Dementia risk prediction

This section summarizes the main findings and discussion points from Study 1 (Reinke et al. 2023)

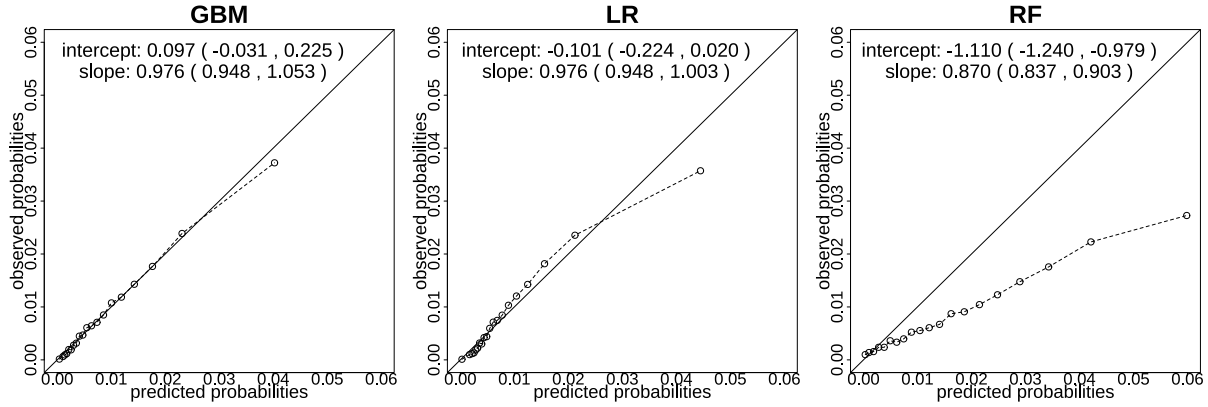
Study cohort

At baseline, the mean age was 74.8 years (SD = 6.6), with a mean age at dementia diagnosis of 82.3 years (SD = 6.3). 38% of the participants were male. The mean time of follow-up was 17.8 quarters (SD = 11.3).

Main Results

LR exhibited the best discriminatory power (C-index = 0.714, 95% CI = 0.708–0.720), closely followed by GBM (0.707, 0.700–0.713). The pairwise differences of the C-index were found to be significantly different from zero (mean = 0.007, 0.005–0.009). The random forest model exhibited a markedly lower discriminatory power, as indicated by a C-index of 0.636 (0.628–0.643). LR exhibited the lowest integrated prediction error (0.044, 0.044–0.045), followed by GBM (0.046, 0.046–0.047) and RF (0.105, = 0.104–0.107). Calibration plots (shown in Figure 5) indicated that GBM exhibited strong calibration (intercept = 0.097, slope = 1.024), while LR demonstrated slightly less calibration (−0.101, 0.976). In contrast, RF exhibited significant calibration issues (−1.110, 0.870).

Figure 5: Calibration on test data. Intercepts and slopes calculated by logistic calibration

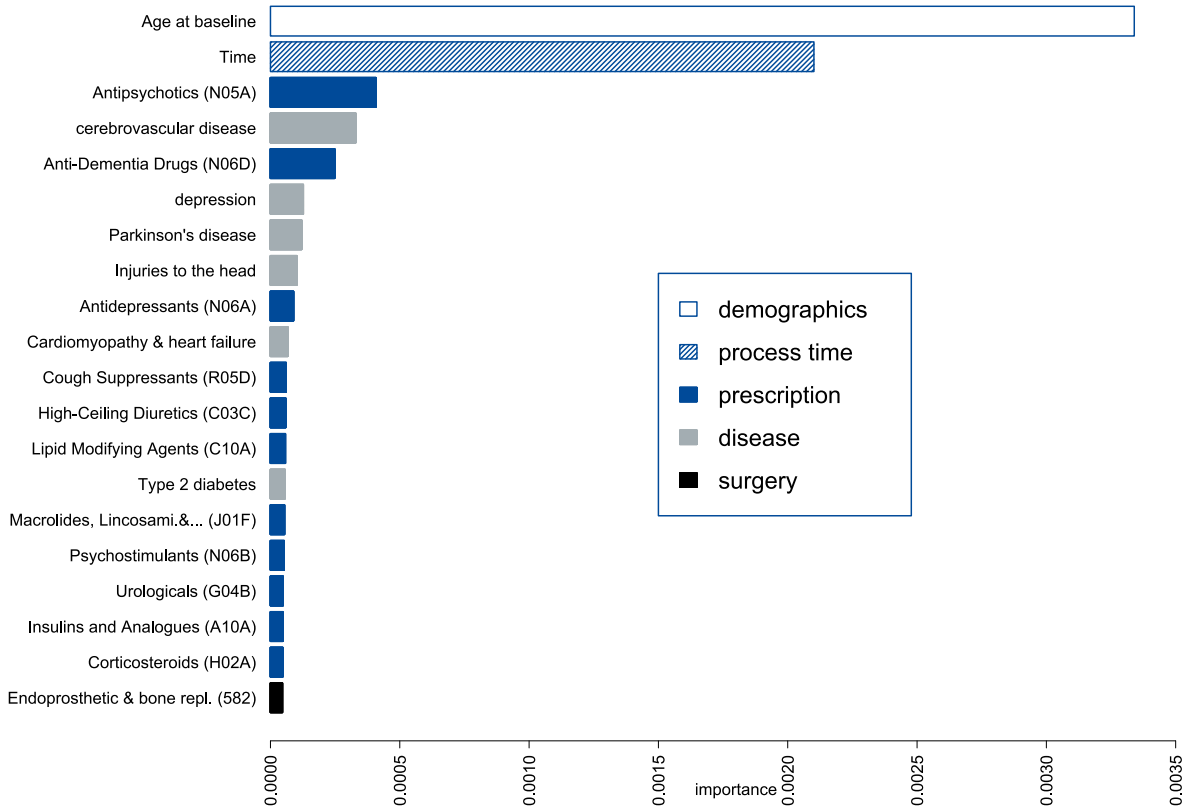


Source: AOK-data 2004-2015, own calculations

Looking at the variable importance of the GBM model (Figure 6), the age at baseline was the most important predictor. Under the medical prescriptions, antipsychotics (ATC-code: N05A), anti-dementia drugs (N06D), and antidepressants (N06A) appear as most influential predictors. Some drugs are associated with conditions that increase the risk of dementia, such as insulin (diabetes) and diuretics (hypertension). Others, such as the group of macrolides, lincosamides and streptogramins (J01F), show a previously unknown relationship to the risk of dementia. Moreover, well-established features such

as cerebrovascular disease, depression and Parkinson's disease have also been identified as important contributors. The odds ratios of the LR are reflect a similar order in relation to variable importance (shown in Figure 7). The odds ratios indicated that the group of macrolides, lincosamides and streptogramins is associated with a reduced risk of dementia.

Figure 6: Variable importance top 20 from the GBM model.

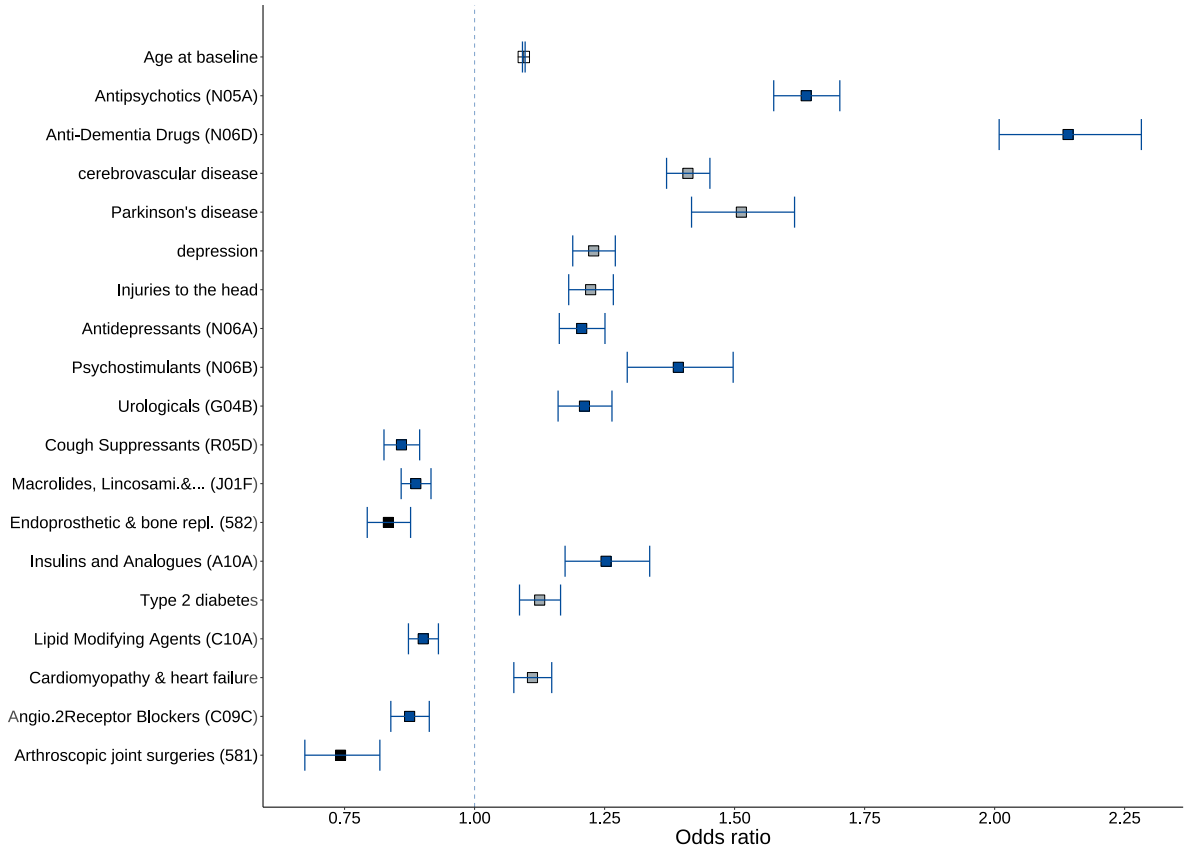


Source: AOK-data 2004-2015, own calculations

Discussion

Both logistic regression and GBM indicated accurate discriminatory power and the results are comparable to or partly better than the results of other studies using claims data (Fukunishi et al. 2020; Nori et al. 2019; Albrecht et al. 2018). The ML methods did not outperform classical regression. However, GBM demonstrated more accurate calibration. The results from regression models are often more straightforward to interpret than those from ML methods. Conversely, variable importance measures provide greater insight into the predictive ability. Among the top 20 predictors, age was identified as the most important one, followed by other established risk factors, including diabetes and late-life depression. Of particular interest was the appearance of a less-established feature, namely the antibacterial medical prescriptions macrolides, lincosamides, and streptogramins, which also emerged as an important predictor.

Figure 7: Odds ratios from logistic regression according to the highest absolute z values.



Source: AOK-data 2004-2015, own calculations

It can be concluded that the utilization of routinely collected claims data provides a valuable supplementary measure for dementia risk detection; specifically, to implement decision support for early dementia screening in a cost-effective manner. The combination of claims data with data-driven methods may serve as a valuable foundation for the identification of previously unknown features associated with the risk of dementia.

5.2. Study 2: The effect of diabetes between education and cognitive function

The following section summarizes the main findings and discussion points from Study 2 (Reinke 2024).

Study cohort

The research sample comprised 26,131 individuals, of whom 18,486 had received a low or middle level of education, while 7,645 had obtained a higher level of education. A total of 1,449 individuals were identified as having diabetes. The individuals were segregated into 14,503 females and 11,628 males. The mean age-standardized reaction time for high educated individuals (-0.2175) was lower (which implying better cognitive performance) than for individuals with low or middle education (0.0895). Non-diabetic individuals demonstrated better cognitive function than diabetic individuals.

Mediation analysis

Regression analyses showed (Table 2) that higher education was associated with a lower risk of diabetes (b=0.1976, p = 0.005). Higher educated individuals have better cognitive function than those with low or middle education (-0.2023, p < 0.001). Individuals with diabetes have a significantly worse cognitive function than individuals without diabetes (0.0617, p = 0.008). The average mediation effect of education through diabetes was -0.00061 and the direct effect of education was -0.20247. This results in a total effect of -0.20307. This implies that individuals with higher educations were, on average, 0.20307 standard deviations faster (demonstrating better cognitive function) than the average for their respective age groups. The indirect effect contributes less than 1% to the total effect.

Table 2: Results of regression models for the mediator and the outcome variable

	Model 1* (mediator)	**Model 2** (outcome)**
dependent variable:	diabetes status	cognitive function
model type:	logistic	OLS
	Reg. coef. (95% CI)	**Reg coef. (95% CI)**
High education (Ref.: low-middle)	-0.1976 (-0.3354; -0.0597)	-0.2023 (-0.2246; -0.1798)
Diabetes (Ref.: No-Diabetes)	-	0.0617 (0.0162; 0.1072)
Number of observations	26,131	26,131
ACME of education (indirect effect)	-0.00061 (-0.00142; -0.00011)	
direct effect of education	-0.20247 (-0.22516; -0.18052)	

regression coefficients and 95% confidents intervals and direct & indirect effect of education from the causal mediation analysis, source: Lifelines data 2006-2015, own calculation.

** Model controlled for: Age, sex, physical activity, obesity, smoking history, income, and hypertension*

*** Model controlled for: age, sex, physical activity, obesity, smoking history, income, comorbidities, and cognition test accuracy*

Discussion

Higher education showed positive effects on cognitive function as well as a lower risk of diabetes. The identification of a significant indirect effect of education on cognitive function via diabetes was of particular interest. These finding suggest that diabetes may act as a potential mediator in the relationship between education and cognitive function.

The findings challenge the conclusions of a previous study (Nakahori et al. 2018), which did not identify a significant association between education and the risk of diabetes. Nevertheless, this study also failed to demonstrate the well-established associations between education and other risk factors for dementia, such as smoking. Another study assumed an interaction effect between education and diabetes, concluding that the effects of the two risk factors were purely additive. This assumption does not contradict the possibility that the relationship may be, in part, a mediated association (Kowall and Rathmann 2023).

In conclusion, this study highlights the role of diabetes in influencing the association between education and cognitive function. People with lower levels of education are twice as disadvantaged in terms of cognitive function, indirectly through their higher risk of diabetes and directly through their education level. The early detection and treatment of diabetes may be crucial for the preservation of cognitive function, particularly among those with lower levels of education.

5.3. Study 3: Diabetes duration and the risk of dementia

The following section presents the main findings and discussion points emerging from Study 3 (Reinke et al. 2022).

Study cohort

The study sample consisted of 57,613 person-years, with a mean follow-up time of 4.18 years and a mean age of 76.9 years at the incident diabetes. The mean aDCSI was 2.34, and 2,278 individuals exhibited no diabetes-related complications. Dementia incidence decreased for one year after a diabetes diagnosis and increased after four years. Severe diabetes cases showed a higher dementia incidence, those treated with insulin having the highest dementia incidence.

Model results

The linear term d of diabetes duration indicated a decreasing risk (model 1: HR(d) = 0.92; 95%CI = 0.90–0.93;) and the significant quadratic d^2 term modulate an increasing dementia risk for longer duration (model 1: HR(d^2) = 1.002; 95%CI = 1.002–1.003) (shown in Table 3). This interplay yielded a U-shaped risk trajectory for dementia across diabetes duration. The predicted values (shown in Figure 8) indicated a decrease in the dementia risk of 26% after one year (predicted HR(d,d^2) = 0.74; 95%CI = 0.70–0.78). The minimum was reached after 4.75 years (predicted HR(d,d^2) = 0.44; 95%CI = 0.39–0.50). The same pattern was observed in the different treatment groups, with a more pronounced U-shaped curve present in the group of insulin users. A year after the first diabetes diagnosis, the risk of dementia was predicted to be 61% higher among those subjects receiving insulin compared to the group receiving no-ADM (predicted HR(d,d^2,ADM_1) = 1.61; 95%CI = 1.21–2.14). No statistically significant differences were found between the other treatment groups and the control group of no-ADM.

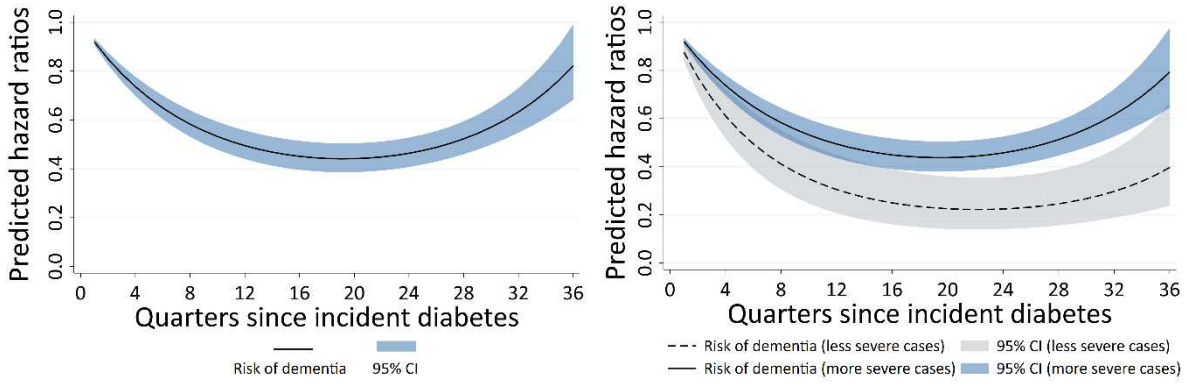
Table 3: Risk of dementia dependent on the duration of diabetes

	Model 1	Model 2	
Variable	**Hazard Ratio (95% CI)**	**Hazard Ratio (95% CI)**	
Time since T2D (d)	$\exp(b_1) = 0.918^{***}$ (0.904 – 0.932)	$\exp(b_1) = 0.926^{***}$ (0.910 – 0.943)	
Time since T2D² (d²)	$\exp(b_2) = 1.002^{***}$ (1.002 – 1.003)	$\exp(b_2) = 1.002^{***}$ (1.001 – 1.003)	
Treatment (ADM_i) (ref. ADM_0)			
Insulin (ADM_1)		$\exp(b_{3,1}) = 3.609^{**}$	(2.305 – 5.651)
Oral ADM (ADM_2)		$\exp(b_{3,2}) = 1.168$	(0.926 – 1.471)
Mixed ADM (ADM_3)		$\exp(b_{3,3}) = 2.186^{*}$	(1.000 – 4.775)
Treatment (ADM_i) × Time since T2D (d)			
Insulin (d, ADM_1)		$\exp(b_{4,1}) = 0.862^{***}$	(0.794 – 0.937)
Oral ADM (d, ADM_2)		$\exp(b_{4,2}) = 0.981$	(0.945 – 1.018)
Mixed ADM (d, ADM_3)		$\exp(b_{4,3}) = 0.966$	(0.849 – 1.100)
Treatment (ADM_i) × Time since T2D² (d^2)			
Insulin (d^2, ADM_1)		$\exp(b_{5,1}) = 1.004^{***}$	(1.001 – 1.006)
Oral ADM (d^2, ADM_2)		$\exp(b_{5,2}) = 1.000$	(0.999 – 1.002)
Mixed ADM (d^2, ADM_3)		$\exp(b_{5,3}) = 0.999$	(0.995 – 1.004)

**P value <0.10. **P value <0.05. ***P value <0.01. All models controlled for: age, sex, comorbidity and aDCSI. Model 1 explore the total duration effect. Models 4 explore the duration effect by treatment groups. Source: AOK-data 2004-2015, own calculation.*

The U-shaped trend was observed across both age categories (<85 and 85+) and all treatment groups, except for mixed ADM. The stratification of diabetes complications has elevated the U-shaped curve for severe cases, indicating that the decrease in dementia risk is stronger in less severe diabetes cases.

Figure 8: Predicted risk of dementia over time since diabetes.



Source: AOK-data 2004–2015, own calculations.

Discussion

The risk of dementia exhibited a U-shaped trajectory over the duration of diabetes. The underlying cause of this trajectory may be multifaceted. Older adults have greater exposure to the medical

system, resulting in higher rates of dementia diagnosis immediately after a diagnosis of diabetes (Kulzer et al. 2013; Müters 2013). This may indicate a monitoring effect. Early diabetes therapy may delay cognitive decline (McIntosh and Nation 2019; Yaffe et al. 2012), but compliance and treatment complexity may reduce this effect (Shrivastava et al. 2013). The early symptoms of dementia are often unnoticed, leading to a delayed diagnosis. Furthermore, cognitive decline may make diabetes management more difficult, increasing the risk of hypo- or hyperglycemic conditions (Puttanna and Padinjakara 2017; Hopkins et al. 2016). Microvascular and macrovascular complications in diabetes may reduce the benefits of therapy over time. This assumption is supported by the less pronounced U-shape in severe diabetes cases at baseline. The increase in dementia risk five or more years after an incidence diabetes diagnosis probably reflects the long-term effects of diabetes on cognition.

The findings are in accordance with the findings of previous studies, which indicate a link between the progression of diabetes and an increased risk of dementia (Chiu et al. 2015). Nevertheless, our findings indicate that the duration of diabetes is an independent risk factor, irrespective of the severity or progression of the disease. Our results differ from those reported by Wu et al (Wu et al. 2015), who found an association between diabetes and increased risk of cognitive decline only in prevalent, not in incident diabetes.

The conclusions of this study suggest that physicians should maintain cognitive monitoring in diabetic patients beyond two years post-diagnosis, considering treatment outcomes and treatment adherence. Future research should investigate whether the prevention, detection, and treatment of diabetes may also contribute to a reduction in the risk of dementia.

5.4. Study 4: Late-life depression and dementia

This section summarizes the main findings and discussion points from Study 4 (Heser et al. 2020).

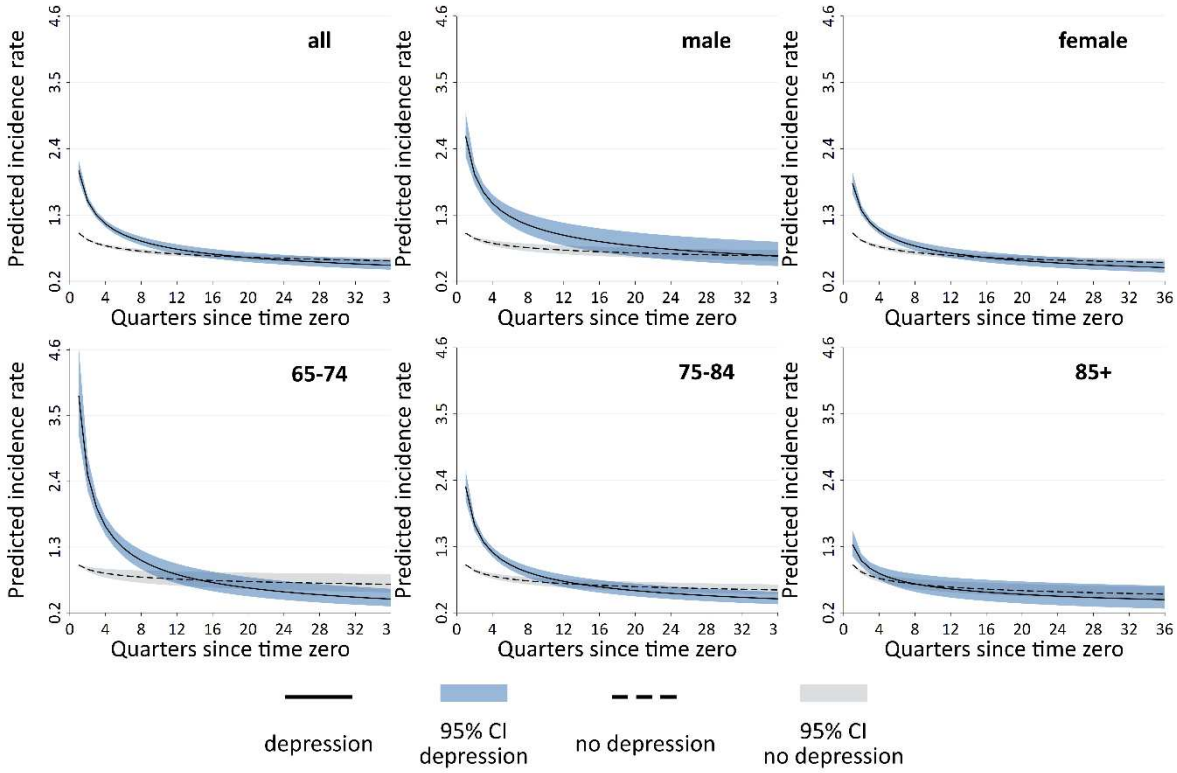
Study cohort

The analyzed sample consisted of 97,110 individuals with mean follow-up period of 5.82 years. The mean age at baseline was 74.7 years (SD = 6.6) for individuals with depression, 78.1 years (SD = 6.1) without depression. A total of 12,668 individuals were identified with incident depression, of whom 2,584 subsequently developed dementia. The incidence of dementia was found to be approximately double that observed in individuals without depression, with an increase in risk corresponding to the severity of depression. The incidence of dementia was slightly higher in females and increased with advancing age.

Model results

Stratified analysis showed that incident depression increased the subsequent risk of dementia for all age groups and genders. The temporal association of depression and dementia was modeled by an interaction effect of depression and the (logarithmic) analysis time. The charts in Figure 9 show the predicted values for the total interaction effect concerning the dementia risk of depressed and non-depressed individuals for each time interval of the analysis time. In the initial quarter, following the incident of depression, the risk of developing dementia was 2.04 times higher in individuals with depression than in those without depression. This increased risk diminished over time. By the 11th quarter, the dementia risk was statistically not different between the two groups. After 20 quarters, both cohorts exhibited an equivalent dementia risk. Depressed men had a 2.61-fold higher risk of dementia in the first quarter and no longer differed significantly from non-depressed men after 13 quarters. Depressed women had a 1.82-fold higher risk of dementia at the beginning and no longer differed significantly after 8 quarters. Depressed individuals in the 65–74 age group exhibited a 3.84-fold increased risk of developing dementia during the initial quarter, with this elevated risk persisting for nine quarters. Individuals aged 75–84 years with depression exhibited a 2.3-fold elevated risk of dementia, which persisted for eight quarters. The elevated risk of dementia for individuals aged 85 years with depression was observed for a period of three quarters, with an initial risk increase of 1.34-fold.

Figure 9: Predicted incidence rates for dementia risk of persons with and without depression dependent on time.



Source: AOK-data 2004–2015, own calculations.

Discussion

Incident late-life depression is associated with an increased risk of dementia within the subsequent three years; however, the association diminishes thereafter. The temporal association was particularly strong between the ages of 65 and 74 and became weaker with advancing age. These findings support the assumption that late-life depression can be a prodrome of dementia, which does not exclude the possibility of mid-life depression being a risk factor.

Prior research has indicated that the strength of the association between depression and subsequent risk of Alzheimer's disease declined when the time period between both diagnoses was extended from zero to seven years (Tapiainen et al. 2017). Another study identified an increase in depressive symptoms as early as a decade prior to dementia, with a threshold of approximately four years before dementia diagnosis (Singh-Manoux et al. 2017). A further study found a higher dementia risk only in increasing trajectories of depressive symptoms (Mirza et al. 2016).

It can be concluded that late-life depression can be a prodrome of dementia and signal an elevated risk of dementia in the near future. A diagnosis of late-life depression should prompt diagnostic

examination and monitoring for cognitive and functional impairment, particularly in the population between 65 and 74 years of age. The potential of immediate and successful treatment of late-life depression to reduce the risk of dementia thus represents an important area for future research.

6. Discussion

This chapter summarizes the findings from the underlying studies of this thesis, contextualizing them within the overarching objective, the theoretical framework and, especially, with open points that need to be discussed. Subsequently, the strengths and limitations of the studies are discussed, with a particular focus on the data sources used. Following this, an outlook and perspectives on future research will be discussed.

The objective of this thesis was to provide evidence for the support for the early detection and prevention of dementia by identifying and obtaining a better understanding of the dementia risk factors by examining their trajectories and pathways.

The key findings of the four underlying studies are summarized below and their new contributions to dementia research are highlighted:

- Study 1 was the first study using German health claims data for dementia risk predictions. And a new, potentially protective, factor was detected as an important predictor for dementia.
- Study 2 was the first study to demonstrate that a portion of the effect of education on cognitive function is mediated by diabetes.
- Study 3 was the first study, that identified a U-shaped dementia risk trajectory for the association to diabetes duration.
- Study 4 was the first study that has shown significant differences in the risk trajectory between late-life depression and subsequent dementia in different age groups and for both sexes.

In light of these findings, some open points require further discussion. These include the transferability of the findings, approaches and underlying mechanisms into further contexts and with other outcomes. The combination of health claims data and data-driven prediction methods represents a promising approach to the identification of risk and protective factors as well as the discovery of potential new factors (Thesmar et al. 2019). A life course approach is required for the early detection and prevention of dementia. However, it remains to be determined whether the results can be transferred to groups outside the analyzed data set. Addressing this question is crucial to the effective implementation of this approach in real-world scenarios. One potential approach is to conduct an evaluation using an external data source, such as the BfArM data referenced in the outlook and perspectives section later in this chapter. In light of these findings, it seems reasonable to question whether the combination of data and methods employed, which has demonstrated efficacy in this context, is applicable to addressing equivalent questions in other contexts or with different outcomes. It was concluded that late-life depression does not act as a risk factor for dementia; rather, it seems to be a prodrome of dementia. Consequently, depressive symptoms in late life may serve as indicators

for early detection of dementia. These findings illustrate that valuable insights can be gained by examining whether a relationship can be characterized by a mechanism other than risk or protection. Accordingly, further links and pathways of factors associated with dementia may be subjected to further investigation. One relationship that has already been called into question is that between diabetes and dementia. Previous studies have proposed the hypothesis that Alzheimer's disease may be regarded as a form of type 3 diabetes, given the significant overlap in the underlying pathology between these conditions (Rorbach-Dolata and Piwowar 2019; La Monte et al. 2018; Kandimalla et al. 2017; Steen et al. 2005). The ongoing process of questioning the evidence regarding the nature of associations is exemplified by the fact that the Lancet Commission classified both depression and diabetes as late-life risk factors in 2020 (Livingston et al. 2020) and concluded in the most recent report of 2024 that there is only sufficient evidence for both factors to indicate a risk association in mid-life (Livingston et al. 2024). This again underscores the necessity of a life course perspective for the identification and evaluation of risk factors.

The findings of a partial mediating effect of diabetes on the pathway between education and cognitive function suggests that the prevention of diabetes may be particularly beneficial for people with low education, as it may have potentially greater effects on reducing the risk of dementia than for those with higher education. The interrelationship between a protective factor from early life (education) and a risk factor from mid-life (diabetes) emphasizes the significance of adopting a life course perspective in the investigation of dementia risk factors. Nevertheless, the contribution of diabetes to the total effect of education on cognitive function was found to represent a relatively minor proportion. This prompts the question of what other factors or mechanisms may be involved in the pathway between education and cognitive function or in the development of dementia. As previously stated, the theoretical perspective of cognitive reserve offers a comprehensive and well-established theory for understanding this association. The theory is based on the utilization of coping strategies against the consequences of brain degeneration through cognitive compensatory processes. Education, in this context, represents a significant dimension in this theoretical construct (Nogueira et al. 2022; Peña-González et al. 2020). The evidence for the protective effect of a higher cognitive reserve is strong (Meng and D'Arcy 2012; Liu et al. 2024), yet the mechanisms behind cognitive reserve remain unclear. In order to implement intervention strategies that can utilize or implement this protective effect, it is necessary to explore the underlying mechanisms of cognitive reserve (Pappalettera et al. 2024), which includes the dimensions of education.

Another issue to be clarified is the identified U-shaped pattern of dementia risk following a diabetes diagnosis. It provides a closer insight into the trajectory of dementia risk and might indicate phases in the duration of diabetes that are particularly sensitive to the risk of dementia. Nevertheless, the

mechanism through which the risk of dementia decreases after a diagnosis of diabetes remains unclear. One potential explanation for this observed decrease in risk may be partly attributed to a monitoring effect, whereby individuals with multiple conditions or diagnoses (such as diabetes and dementia) are more likely to seek healthcare services, which leads to detection and diagnosis of these conditions concurrently. Consequently, the number of cases of dementia identified in the period following a diabetes diagnosis is reduced, due to the fact that the initial diagnosis was made earlier than would otherwise have been the case. It can be assumed that contact with the health system results in an increased awareness and detection rate of dementia. Consequently, the initial decrease may be attributed to the nature of the data or the data-generating process rather than a disease-associated process. Nevertheless, as the decrease in the dementia risk may be a protective mechanism, it is essential to clarify and investigate this mechanism in order to identify potential intervention strategies. This should be addressed in future research.

Strengths & limitations of the underlying studies

The four underlying studies of this thesis exhibit a number of overlapping strengths and weaknesses, which predominantly relate to the data basis used or the associated study design. For both data sources large population-based samples with 250,000 (AOK-data) respectively 167,729 individuals (Lifelines) were available and have been analyzed. The AOK-data is representative for the German population aged 50+ in terms of morbidity and mortality (as illustrated in Figure 3 (Fink 2017)) and the Lifelines data covers about 10% of the population of the northern Netherlands. The AOK-data also includes individuals living in nursing homes which are often very morbid and therefore typically not covered by health surveys. It may be inferred that the Lifelines data set is not fully representative of the population with severe morbidity. This is indicated by a diabetes prevalence of 5.7% for the study sample (age 50+) compared to 7.5% in the Netherland population in the age-group 20–79 years (Aguirre et al. 2013). This suggests the suspicion of a selection bias, which is a typical problem in survey data (Elston 2021; Strandhagen et al. 2010). Furthermore, the sample of the AOK-data had a follow-up of up to 11 years with quarterly information intervals, which enabled the setup of time-varying covariates. On the other hand, the used Lifelines sample had a varying follow-up time of also up to 11 years (2007-2017), however only two measurement points and the period between the baseline (2007-2013) and second assessment (2014-2017). This is a quite short period, given the slow progression of diabetes and cognitive decline analyzed here. A further issue that arises in the context of survey data is that of attrition (Saiepour et al. 2019), which is not relevant in the context of claims data. But almost 30% of the participants from the baseline assessment did not participate at the second one (Sijtsma et al. 2022), related to the complete population of the Lifelines study. The survey data from the Lifelines Study contains individual laboratory values and test results, including the Cogstate examination and

the HbA1c value. This enables broader definitions of diabetes to be made, which also include undiagnosed cases. However, diagnoses that cannot be replaced by individual tests are typically only available in survey data as self-reported information, which may be subject to recall bias. The AOK-data set does not include any laboratory test results or other parameters; instead, it provides diagnoses from both inpatient and outpatient settings, which have been made by healthcare professionals. It should be noted, however, that the diagnoses only occurred in the data when the insured individuals interacted with the healthcare system. Consequently, diagnoses can only be identified for individuals who went to a doctor, which may result in lower incidence and prevalence estimates due to a potential reporting bias. In contrast, previous research has demonstrated that the age-specific prevalence and incidence of dementia in AOK-data are comparable to those observed in the entirety of the German population (Doblhammer et al. 2015). A further issue is that diagnoses in claims data indicate only the time of diagnosis rather than the time of onset. This is an important detail to consider in the context of slow progressive diseases such as dementia or diabetes, as well as when analyzing trajectories over time. Moreover, the data collected on medical prescriptions is limited to the specific drug prescribed, without any information regarding the actual intake. A key strength of the Lifelines data is the inclusion of information pertaining to lifestyle factors, health behaviors and socioeconomic factors such as education and income, which are not part of health claims data. This is of special importance because the proportion of persons with low socioeconomic status is higher in the AOK than in other statutory health insurance companies and also in comparison with private health insurance system (Epping et al. 2021). Neither the AOK-data nor the Lifelines data contains genetic information that plays a role for the dementia risk (van der Lee et al. 2018; Loy et al. 2014). The AOK-data underlies issues of truncation and censoring (Weißbach et al. 2024; Weißbach et al. 2021). Ignoring these issues can lead to biased estimates. The problems of right censoring and left truncation were taken into account in the study designs (Cain et al. 2011). However, on the one hand, ignoring left-censoring may result in an underestimation of disease rates (Weißbach et al. 2021). On the other hand, if it is assumed that exposure and outcome rates are equally affected, this should not affect the mechanism between exposure and outcome.

In all studies, a number of sensitivity analyses were conducted to ensure that the findings were not dependent on the model strategies employed and to evaluate the robustness of the study designs. Moreover, the outcomes and exposures, including dementia, depression, and diabetes, as well as the measure for diabetes severity (aDCSI), were validation by established strategies in studies that used the AOK-data. With regard to Study 1, it is important to note that the prediction models were validated using data from the same source. However, it would be preferable to conduct a validation process using data from an external source. It should be noted for Study 2 that the Cogstate Brief Battery was

not measured at baseline, thus not allowing for an examination of changes in cognitive function within the context of the analyses.

Outlook and perspectives

In light of the aforementioned limitations, the question arises as to which data are necessary for future research on dementia risk factors. In a hypothetical scenario, large population-based data would be available covering the entire life-course from the time of birth (or even gestation) until the end of life. This data would include medical history, including laboratory parameters, demographic information, socio-economic information, geographical information (for linkage with environmental data) and genetic information. Initially, the availability of such data seems cost-intensive and difficult to realize. With the exception of genetic information, all of this data is already existing for the majority of the German population; without a survey, but in the form of routinely collected data. The BfArM has already started to provide data on all individuals with statutory health insurance in Germany (Forschungsdatenzentrum Gesundheit 2024). In content and structure, this data corresponds to the AOK-data described in this thesis and contains a wealth of information on medical history and almost all information on the utilization of healthcare services, as well as individual demographic and geographical information. Following an extensive period of public and political discourse, the electronic health record (EHR) and the prospect of its utilization for research purposes have been initiated (Bundesministerium der Justiz Bundesministerium der Justiz 2024b, 2024a). This has enabled the incorporation of a range of information into the EHR, including laboratory data, clinical assessments and medical correspondence. Some socio-economic information is also still available by the health insurance providers or could be accessed via the statutory pension fund. As health insurance is mandatory in Germany, almost all people living in the country are covered by this data. And even if the BfArM data only refers to people in the statutory health insurance system, it still covers almost 90% of the German population. Accordingly, it may be feasible to get closer to the theoretical 45 % of preventable dementia. And by discovering new or better understandings of risk factors, it may be possible to further reduce the currently non-modifiable part of the dementia risk, which was estimated at 60% in 2020 and fell to 55% in 2024.

However, the political implementation of such a project (in particular the linking of data) and crucial issues of individual data protection are not respected in these considerations. The availability of such data in the near future will probably only remain a hypothetical scenario. Nevertheless, the data provided by the BfArM represents a promising advance for future research on dementia risk factors, offering the opportunity to employ a life-course approach. A further positive example is the combination of routinely collected data with information from other data sources. A few studies used data from the UK Biobank, a large biomedical database containing information on genetics, lifestyle

and biological samples and more from the UK (Sudlow et al. 2015), and linked this data with hospital admission and primary care data (Wilkinson et al. 2019) as well as EHRs from the National Health Service (NHS) in the UK (Parra et al. 2022).

Existing public health strategies against dementia demonstrate that some prevention and early detection measures have already been addressed or implemented. For example, both the National Plan to Address Alzheimer's Disease from the USA and the “Nationale Demenzstrategie” (National Dementia Strategy) from Germany highlight the necessity for further research into modifiable risk factors, given the current limited evidence base for some potential factors. To this end, efforts should be made to improve the availability and accessibility of data (HHS 2023; bmfsfj 2020).

7. Conclusion

In consideration of the ageing world population and thus the drastically increasing number of dementia cases in the future, dementia prevention is urgently needed, possible and, last but not least, as this work also demonstrates, available.

Ideally, dementia should be either prevented or the risk of dementia reduced. The key to achieving this objective is to tackle the risk factors associated with dementia. Cases that have not been prevented should be recognized as early as possible so that the progression of dementia and the associated limitations can be delayed. These issues do not only affect the older population or the late-life phase, but are a challenge across the entire life course. The objective of this thesis was to identify and explain risk factors and mechanisms associated with dementia as well as to provide evidence to support the early detection of this condition. Therefore, perspectives and methods from the fields of demography, epidemiology and public health research were applied in four studies, leading to the following conclusion.

Early detection is particularly important against the background of the latest developments in dementia therapy. The combination of routinely collected data and data-driven approaches should be applied for a cost-effective tool of early dementia detection. Additionally, late-life depression may signal an increased risk of dementia in the near future and should be used for early detection by a prompt comprehensive monitoring for cognitive and functional impairment, particularly in individuals between 65-74 years of age. The risk factor of diabetes and the protective factor of education represent two modifiable factors that should be targeted for the prevention of dementia. The initial decrease in the dementia risk after a diabetes diagnosis suggests potential benefits from diabetes treatment interventions. The cognitive functions of individuals with diabetes should be closely monitored, especially beyond two years after the diagnosis of diabetes. Detection and treatment compliance of diabetes might also be beneficial regarding the development of dementia. The vulnerable group of lower-educated individuals can particularly benefit, i.e. maintain good cognitive function and prevent dementia.

Future research and identification of dementia risk factors and their mechanisms should be focused. This thesis and the underlying studies contribute to the fight against dementia by drawing implications and providing a basis for further research.

8. References

Abolhasani, Ehsan; Hachinski, Vladimir; Ghazaleh, Nargess; Azarpazhooh, Mahmoud Reza; Mokhber, Naghmeh; Martin, Janet (2023): Air Pollution and Incidence of Dementia: A Systematic Review and Meta-analysis. In *Neurology* 100 (2), e242-e254. DOI: 10.1212/WNL.0000000000201419.

Aguirre, F.; Brown, A.; Cho, N. H.; Dahlquist, G.; Dodd, S.; Dunning, T. et al. (2013): IDF Diabetes Atlas. Sixth edition. (Sixth ed.): International Diabetes Federation.

Aisen, Paul S.; Cummings, Jeffrey; Jack, Clifford R.; Morris, John C.; Sperling, Reisa; Frölich, Lutz et al. (2017): On the path to 2025: understanding the Alzheimer's disease continuum. In *Alzheimer's research & therapy* 9 (1), p. 60. DOI: 10.1186/s13195-017-0283-5.

Albrecht, Jennifer S.; Hanna, Maya; Kim, Dure; Perfetto, Eleanor M. (2018): Predicting Diagnosis of Alzheimer's Disease and Related Dementias Using Administrative Claims. In *Journal of managed care & specialty pharmacy* 24 (11), pp. 1138–1145. DOI: 10.18553/jmcp.2018.24.11.1138.

Almeida, O. P.; Hankey, G. J.; Yeap, B. B.; Golledge, J.; Flicker, L. (2017): Depression as a modifiable factor to decrease the risk of dementia. In *Translational psychiatry* 7 (5), e1117. DOI: 10.1038/tp.2017.90.

American Psychiatric Association (2013): Diagnostic and statistical manual of mental disorders. DSM-5. 5. edition. Washington, DC: American Psychiatric Publishing. Available online at https://www.lehmanns.de/media/74773348.

Aretz, Benjamin; Janssen, Fanny; Vonk, Judith M.; Heneka, Michael T.; Boezen, H. Marike; Doblhammer, Gabriele (2021): Long-term exposure to fine particulate matter, lung function and cognitive performance: A prospective Dutch cohort study on the underlying routes. In *Environmental research* 201, p. 111533. DOI: 10.1016/j.envres.2021.111533.

Arevalo-Rodriguez, Ingrid; Smailagic, Nadja; Roqué I Figuls, Marta; Ciapponi, Agustín; Sanchez-Perez, Erick; Giannakou, Antri et al. (2015): Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). In *The Cochrane database of systematic reviews* 2015 (3), CD010783. DOI: 10.1002/14651858.CD010783.pub2.

Arvanitakis, Zoe; Shah, Raj C.; Bennett, David A. (2019): Diagnosis and Management of Dementia: Review. In *JAMA* 322 (16), pp. 1589–1599. DOI: 10.1001/jama.2019.4782.

Avgerinos, Konstantinos I.; Ferrucci, Luigi; Kapogiannis, Dimitrios (2021): Effects of monoclonal antibodies against amyloid-β on clinical and biomarker outcomes and adverse event risks: A systematic review and meta-analysis of phase III RCTs in Alzheimer's disease. In *Ageing research reviews* 68, p. 101339. DOI: 10.1016/j.arr.2021.101339.

Bäckman, Lars; Wahlin, Åke; Small, Brent J.; Herlitz, Agneta; Winblad, Bengt; Fratiglioni, Laura (2004): Cognitive Functioning in Aging and Dementia: The Kungsholmen Project. In *Aging, Neuropsychology, and Cognition* 11 (2-3), pp. 212–244. DOI: 10.1080/13825580490511099.

Bai, Wei; Chen, Pan; Cai, Hong; Zhang, Qinge; Su, Zhaohui; Cheung, Teris et al. (2022): Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. In *Age and ageing* 51 (8). DOI: 10.1093/ageing/afac173.

Barbiellini Amidei, Claudio; Fayosse, Aurore; Dumurgier, Julien; Machado-Fragua, Marcos D.; Tabak, Adam G.; van Sloten, Thomas et al. (2021): Association Between Age at Diabetes Onset and Subsequent Risk of Dementia. In *JAMA* 325 (16), pp. 1640–1649. DOI: 10.1001/jama.2021.4001.

Bellou, Vanesa; Belbasis, Lazaros; Tzoulaki, Ioanna; Evangelou, Evangelos (2018): Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses. In *PloS one* 13 (3), e0194127. DOI: 10.1371/journal.pone.0194127.

Ben Miled, Zina; Haas, Kyle; Black, Christopher M.; Khandker, Rezaul Karim; Chandrasekaran, Vasu; Lipton, Richard; Boustani, Malaz A. (2020): Predicting dementia with routine care EMR data. In *Artificial intelligence in medicine* 102, p. 101771. DOI: 10.1016/j.artmed.2019.101771.

Ben-Shlomo, Y.; Kuh, D. (2002): A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. In *International Journal of Epidemiology* 31 (2), pp. 285–293. DOI: 10.1093/ije/31.2.285.

Biessels, Geert Jan; Staekenborg, Salka; Brunner, Eric; Brayne, Carol; Scheltens, Philip (2006): Risk of dementia in diabetes mellitus: a systematic review. In *The Lancet. Neurology* 5 (1), pp. 64–74. DOI: 10.1016/S1474-4422(05)70284-2.

Blotenberg, Iris; Boekholt, Melanie; Michalowsky, Bernhard; Platen, Moritz; Rodriguez, Francisca S.; Teipel, Stefan et al. (2024): What influences life expectancy in people with dementia? Social support as an emerging protective factor. In *Age and ageing* 53 (3). DOI: 10.1093/ageing/afae044.

Blotenberg, Iris; Hoffmann, Wolfgang; Thyrian, Jochen René (2023): Dementia in Germany: Epidemiology and Prevention Potential. In *Deutsches Arzteblatt international* 120 (27-28), pp. 470–476. DOI: 10.3238/arztebl.m2023.0100.

bmfsfj (Ed.) (2020): Nationale Demenzstrategie. 1. Auflage. Bundesministerium für Familie, Senioren, Frauen und Jugend. Available online at https://www.nationale-demenzstrategie.de/fileadmin/nds/pdf/2020-07-01_Nationale_Demenzsstrategie.pdf.

Breiman, Leo (2001a): Random Forests. In *Machine Learning* 45 (1), pp. 5–32. DOI: 10.1023/A:1010933404324.

Breiman, Leo (2001b): Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). In *Statist. Sci.* 16 (3). DOI: 10.1214/ss/1009213726.

Brett, Benjamin L.; Gardner, Raquel C.; Godbout, Jonathan; Dams-O'Connor, Kristen; Keene, C. Dirk (2022): Traumatic Brain Injury and Risk of Neurodegenerative Disorder. In *Biological psychiatry* 91 (5), pp. 498–507. DOI: 10.1016/j.biopsych.2021.05.025.

Bundesministerium der Justiz Bundesministerium der Justiz (2024a): Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz – DigiG). DigiG. In *Bundesgesetzblatt* 2024 (101). Available online at https://www.bundesgesundheitsministerium.de/service/gesetze-und-verordnungen/detail/digital-gesetz.html, checked on 11/15/2024.

Bundesministerium der Justiz Bundesministerium der Justiz (2024b): Gesetz zur verbesserten Nutzung von Gesundheitsdaten. GDNG. In *Bundesgesetzblatt* 2024 (102). Available online at https://www.bundesgesundheitsministerium.de/service/gesetze-und-verordnungen/detail/gesundheitsdatennutzungsgesetz.html, checked on 11/15/2024.

Butters, Meryl A.; Young, Jeffrey B.; Lopez, Oscar; Aizenstein, Howard J.; Mulsant, Benoit H.; Reynolds, Charles F. et al. (2008): Pathways linking late-life depression to persistent cognitive impairment and dementia. In *Dialogues in clinical neuroscience* 10 (3), pp. 345–357. DOI: 10.31887/DCNS.2008.10.3/mabutters.

Byers, Amy L.; Yaffe, Kristine (2011): Depression and risk of developing dementia. In *Nature reviews. Neurology* 7 (6), pp. 323–331. DOI: 10.1038/nrneurol.2011.60.

Cadarette, Suzanne M.; Wong, Lindsay (2015): An Introduction to Health Care Administrative Data. In *The Canadian journal of hospital pharmacy* 68 (3), pp. 232–237. DOI: 10.4212/cjhp.v68i3.1457.

Cain, Kevin C.; Harlow, Siobán D.; Little, Roderick J.; Nan, Bin; Yosef, Matheos; Taffe, John R.; Elliott, Michael R. (2011): Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. In *American journal of epidemiology* 173 (9), pp. 1078–1084. DOI: 10.1093/aje/kwq481.

Cao, Qing; Tan, Chen-Chen; Xu, Wei; Hu, Hao; Cao, Xi-Peng; Dong, Qiang et al. (2020): The Prevalence of Dementia: A Systematic Review and Meta-Analysis. In *JAD* 73 (3), pp. 1157–1166. DOI: 10.3233/JAD-191092.

Chang, Hsien-Yen; Weiner, Jonathan P.; Richards, Thomas M.; Bleich, Sara N.; Segal, Jodi B. (2012): Validating the adapted Diabetes Complications Severity Index in claims data. In *The American journal of managed care* 18 (11), pp. 721–726.

Chang, Po-Yin; Wang, I-T Ing; Chiang, Chern-En; Chen, Chen-Huan; Yeh, Wan-Yu; Henderson, Victor W. et al. (2021): Vascular complications of diabetes: natural history and corresponding risks of dementia in a national cohort of adults with diabetes. In *Acta diabetologica* 58 (7), pp. 859–867. DOI: 10.1007/s00592-021-01685-y.

Chen, Yi; Tysinger, Bryan; Crimmins, Eileen; Zissimopoulos, Julie M. (2019): Analysis of dementia in the US population using Medicare claims: Insights from linked survey and administrative claims data. In *Alzheimer's & dementia (New York, N. Y.)* 5, pp. 197–207. DOI: 10.1016/j.trci.2019.04.003.

Cherbuin, Nicolas; Kim, Sarang; Anstey, Kaarin J. (2015): Dementia risk estimates associated with measures of depression: a systematic review and meta-analysis. In *BMJ open* 5 (12), e008853. DOI: 10.1136/bmjopen-2015-008853.

Chiu, Wei-Che; Ho, Wen-Chao; Liao, Ding-Lieh; Lin, Meng-Hung; Chiu, Chih-Chiang; Su, Yu-Ping; Chen, Pau-Chung (2015): Progress of Diabetic Severity and Risk of Dementia. In *The Journal of clinical endocrinology and metabolism* 100 (8), pp. 2899–2908. DOI: 10.1210/jc.2015-1677.

Choi, Daein; Choi, Seulggie; Park, Sang Min (2018): Effect of smoking cessation on the risk of dementia: a longitudinal study. In *Annals of clinical and translational neurology* 5 (10), pp. 1192–1199. DOI: 10.1002/acn3.633.

Cleret de Langavant, Laurent; Bayen, Eléonore; Bachoud-Lévi, Anne-Catherine; Yaffe, Kristine (2020): Approximating dementia prevalence in population-based surveys of aging worldwide: An unsupervised machine learning approach. In *Alzheimer's & dementia (New York, N. Y.)* 6 (1), e12074. DOI: 10.1002/trc2.12074.

Coughlin, S. S. (1990): Recall bias in epidemiologic studies. In *Journal of clinical epidemiology* 43 (1), pp. 87–91. DOI: 10.1016/0895-4356(90)90060-3.

Croxson, S. C.; Jagger, C. (1995): Diabetes and cognitive impairment: a community-based study of elderly subjects. In *Age and ageing* 24 (5), pp. 421–424. DOI: 10.1093/ageing/24.5.421.

Dannefer, Dale (1984): Adult Development and Social Theory: A Paradigmatic Reappraisal. In *American Sociological Review* 49 (1), pp. 100–116. Available online at http://www.jstor.org/stable/2095560.

DeFronzo, Ralph A. (2009): Banting Lecture. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. In *Diabetes* 58 (4), pp. 773–795. DOI: 10.2337/db09-9028.

DeFronzo, Ralph A.; Ferrannini, Ele; Groop, Leif; Henry, Robert R.; Herman, William H.; Holst, Jens Juul et al. (2015): Type 2 diabetes mellitus. In *Nature Reviews Disease Primers* 1 (1), p. 15019. DOI: 10.1038/nrdp.2015.19.

Dekhtyar, Serhiy; Wang, Hui-Xin; Scott, Kirk; Goodman, Anna; Koupil, Ilona; Herlitz, Agneta (2015): A Life-Course Study of Cognitive Reserve in Dementia--From Childhood to Old Age. In *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry* 23 (9), pp. 885–896. DOI: 10.1016/j.jagp.2015.02.002.

Denver, Paul; McClean, Paula L. (2018): Distinguishing normal brain aging from the development of Alzheimer's disease: inflammation, insulin signaling and cognition. In *Neural regeneration research* 13 (10), pp. 1719–1730. DOI: 10.4103/1673-5374.238608.

DGN & DGPPN (Ed.) (2023): S3-Leitlinie Demenzen. Version: 4.0. Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde e. V. (DGPPN) & Deutsche Gesellschaft für Neurologie e. V. (DGN). Available online at https://register.awmf.org/de/leitlinien/detail/038-013, checked on 9/26/2024.

Ding, Ding; Zhao, Qianhua; Wu, Wanqing; Xiao, Zhenxu; Liang, Xiaoniu; Luo, Jianfeng; Hong, Zhen (2020): Prevalence and incidence of dementia in an older Chinese population over two decades: The role of education. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 16 (12), pp. 1650–1662. DOI: 10.1002/alz.12159.

Diniz, Breno S.; Butters, Meryl A.; Albert, Steven M.; Dew, Mary Amanda; Reynolds, Charles F. (2013): Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. In *Br J Psychiatry* 202 (5), pp. 329–335. DOI: 10.1192/bjp.bp.112.118307.

Doblhammer, Gabriele; Barth, Alexander (2018): Prevalence of Morbidity at Extreme Old Age in Germany: An Observational Study Using Health Claims Data. In *Journal of the American Geriatrics Society* 66 (7), pp. 1262–1268. DOI: 10.1111/jgs.15460.

Doblhammer, Gabriele; Fink, Anne; Fritze, Thomas (2015): Short-term trends in dementia prevalence in Germany between the years 2007 and 2009. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 11 (3), pp. 291–299. DOI: 10.1016/j.jalz.2014.02.006.

Doblhammer, Gabriele; Fritze, Thomas; Reinke, Constantin; Fink, Anne (2022): Can dementia become the most prevalent disease at the time of death in Germany? Projections up to the year 2060 for the five most important diseases at the time of death. In *Population Ageing* 15 (2), pp. 523–540. DOI: 10.1007/s12062-022-09365-7.

Döring, Andrea; Paul, Friedemann (2010): The German healthcare system. In *The EPMA journal* 1 (4), pp. 535–547. DOI: 10.1007/s13167-010-0060-z.

Elahi, Fanny M.; Miller, Bruce L. (2017): A clinicopathological approach to the diagnosis of dementia. In *Nature reviews. Neurology* 13 (8), pp. 457–476. DOI: 10.1038/nrneurol.2017.96.

Elovainio, Marko; Lahti, Jari; Pirinen, Matti; Pulkki-Råback, Laura; Malmberg, Anni; Lipsanen, Jari et al. (2022): Association of social isolation, loneliness and genetic risk with incidence of dementia: UK Biobank Cohort Study. In *BMJ open* 12 (2), e053936. DOI: 10.1136/bmjopen-2021-053936.

Elser, Holly; Horváth-Puhó, Erzsébet; Gradus, Jaimie L.; Smith, Meghan L.; Lash, Timothy L.; Glymour, M. Maria et al. (2023): Association of Early-, Middle-, and Late-Life Depression With Incident Dementia in a Danish Cohort. In *JAMA neurology* 80 (9), pp. 949–958. DOI: 10.1001/jamaneurol.2023.2309.

Elston, Dirk M. (2021): Participation bias, self-selection bias, and response bias. In *Journal of the American Academy of Dermatology*. DOI: 10.1016/j.jaad.2021.06.025.

EMA (2024): Leqembi recommended for treatment of early Alzheimer's disease. Re-examination concludes that benefits outweigh risks in a restricted patient population. online. Available online at https://www.ema.europa.eu/en/news/leqembi-recommended-treatment-early-alzheimers-disease, checked on 11/16/2024.

Emilsson, Louise; García-Albéniz, Xabier; Logan, Roger W.; Caniglia, Ellen C.; Kalager, Mette; Hernán, Miguel A. (2018): Examining Bias in Studies of Statin Treatment and Survival in Patients With Cancer. In *JAMA oncology* 4 (1), pp. 63–70. DOI: 10.1001/jamaoncol.2017.2752.

Epping, Jelena; Geyer, Siegfried; Eberhard, Sveja; Tetzlaff, Juliane (2021): Völlig unterschiedlich oder doch recht ähnlich? Die soziodemografische Struktur der AOK Niedersachsen im Vergleich zur niedersächsischen und bundesweiten Allgemein- und Erwerbsbevölkerung. In *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))* 83 (S 02), S77-S86. DOI: 10.1055/a-1553-3565.

Evans, Isobel E. M.; Martyr, Anthony; Collins, Rachel; Brayne, Carol; Clare, Linda (2019): Social Isolation and Cognitive Function in Later Life: A Systematic Review and Meta-Analysis. In *JAD* 70 (s1), S119-S144. DOI: 10.3233/JAD-180501.

Fierini, Fabio (2020): Mixed dementia: Neglected clinical entity or nosographic artifice? In *Journal of the neurological sciences* 410, p. 116662. DOI: 10.1016/j.jns.2019.116662.

Fink, A. (2017): Prevention matters: The impact of modifiable risk factors on neurodegenerative diseases and long-term care need in the ageing population of Germany. Rostock.

Fink, Anne (2014): Dementia and long-term care – an analysis based on German health insurance data. In Gabriele Doblhammer (Ed.): Health Among the Elderly in Germany. New Evidence on Disease, Disability and Care Need: Verlag Barbara Budrich (SERIES ON POPULATION STUDIES, 46), pp. 139–156.

Ford, Elizabeth; Greenslade, Nicholas; Paudyal, Priya; Bremner, Stephen; Smith, Helen E.; Banerjee, Sube et al. (2018): Predicting dementia from primary care records: A systematic review and meta-analysis. In *PloS one* 13 (3), e0194735. DOI: 10.1371/journal.pone.0194735.

Forschungsdatenzentrum Gesundheit (2024): Datensatzbeschreibung des Forschungsdatenzentrums Gesundheit. Available online at https://fdz-gesundheit.github.io/datensatzbeschreibung_fdz_gesundheit/, checked on 11/12/2024.

Freitas, Guilherme B. de; Lourenco, Mychael V.; Felice, Fernanda G. de (2020): Protective actions of exercise-related FNDC5/Irisin in memory and Alzheimer's disease. In *Journal of neurochemistry* 155 (6), pp. 602–611. DOI: 10.1111/jnc.15039.

Fukunishi, Hiroaki; Nishiyama, Mitsuki; Luo, Yuan; Kubo, Masahiro; Kobayashi, Yasuki (2020): Alzheimer-type dementia prediction by sparse logistic regression using claim data. In *Computer methods and programs in biomedicine* 196, p. 105582. DOI: 10.1016/j.cmpb.2020.105582.

Gaertner, Beate; Koschollek, Carmen; Grube, Maike M.; Lüdtke, Denise; Fuchs, Judith; Scheidt-Nave, Christa et al. (2019): Including nursing home residents in a general population health survey in Germany. DOI: 10.25646/6313.

Gao, Mingyue; Kuang, Weihong; Qiu, Peiyuan; Wang, Huali; Lv, Xiaozhen; Yang, Min (2017): The time trends of cognitive impairment incidence among older Chinese people in the community: based on

the CLHLS cohorts from 1998 to 2014. In *Age and ageing* 46 (5), pp. 787–793. DOI: 10.1093/ageing/afx038.

Gardner, Raquel C.; Bahorik, Amber; Kornblith, Erica S.; Allen, Isabel Elaine; Plassman, Brenda L.; Yaffe, Kristine (2023): Systematic Review, Meta-Analysis, and Population Attributable Risk of Dementia Associated with Traumatic Brain Injury in Civilians and Veterans. In *Journal of neurotrauma* 40 (7-8), pp. 620–634. DOI: 10.1089/neu.2022.0041.

Gauthier, S.; Rosa-Neto, P.; Morais, JA.; Webster, C. (2021): World Alzheimer Report 2021. Journey through the diagnosis of dementia. London, England: Alzheimer's Disease International. Available online at https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf.

Georges, Daniela; Rakusa, Elena; Holtz, Anna-Victoria; Fink, Anne; Doblhammer, Gabriele (2023): Dementia in Germany: epidemiology, trends and challenges. In *Journal of health monitoring* 8 (3), pp. 30–48. DOI: 10.25646/11667.

GKV-Spitzenverband (Ed.) (2024): Kennzahlen der gesetzlichen Krankenversicherung. SEPTEMBER 2024. With assistance of Stabsbereich Kommunikation. Available online at https://www.gkv-spitzenverband.de/media/grafiken/gkv_kennzahlen/kennzahlen_gkv_2024_q2/20240909_GKV_Kennzahlen_Booklet_Q2-2024_300dpi_barrierefrei.pdf, checked on 10/22/2024.

Goerdten, Jantje; Čukić, Iva; Danso, Samuel O.; Carrière, Isabelle; Muniz-Terrera, Graciela (2019): Statistical methods for dementia risk prediction and recommendations for future work: A systematic review. In *Alzheimer's & dementia (New York, N. Y.)* 5, pp. 563–569. DOI: 10.1016/j.trci.2019.08.001.

Goodman, Richard A.; Lochner, Kimberly A.; Thambisetty, Madhav; Wingo, Thomas S.; Posner, Samuel F.; Ling, Shari M. (2017): Prevalence of dementia subtypes in United States Medicare fee-for-service beneficiaries, 2011-2013. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 13 (1), pp. 28–37. DOI: 10.1016/j.jalz.2016.04.002.

Gottesman, Rebecca F.; Albert, Marilyn S.; Alonso, Alvaro; Coker, Laura H.; Coresh, Josef; Davis, Sonia M. et al. (2017): Associations Between Midlife Vascular Risk Factors and 25-Year Incident Dementia in the Atherosclerosis Risk in Communities (ARIC) Cohort. In *JAMA neurology* 74 (10), pp. 1246–1254. DOI: 10.1001/jamaneurol.2017.1658.

Graham, Neil Sn; Sharp, David J. (2019): Understanding neurodegeneration after traumatic brain injury: from mechanisms to clinical trials in dementia. In *Journal of neurology, neurosurgery, and psychiatry* 90 (11), pp. 1221–1233. DOI: 10.1136/jnnp-2017-317557.

Grasset, Leslie; Brayne, Carol; Joly, Pierre; Jacqmin-Gadda, Hélène; Peres, Karine; Foubert-Samier, Alexandra et al. (2016): Trends in dementia incidence: Evolution over a 10-year period in France. In *Alzheimer's & Dementia* 12 (3), pp. 272–280. DOI: 10.1016/j.jalz.2015.11.001.

Greß, Stefan (2007): Private health insurance in Germany: consequences of a dual system. In *Healthcare policy = Politiques de sante* 3 (2), pp. 29–37.

Griffiths, Timothy D.; Lad, Meher; Kumar, Sukhbinder; Holmes, Emma; McMurray, Bob; Maguire, Eleanor A. et al. (2020): How Can Hearing Loss Cause Dementia? In *Neuron* 108 (3), pp. 401–412. DOI: 10.1016/j.neuron.2020.08.003.

Gu, Dongqing; Ou, Shan; Liu, Guodong (2022): Traumatic Brain Injury and Risk of Dementia and Alzheimer's Disease: A Systematic Review and Meta-Analysis. In *Neuroepidemiology* 56 (1), pp. 4–16. DOI: 10.1159/000520966.

Gueorguieva, Ivelina; Willis, Brian A.; Chua, Laiyi; Chow, Kay; Ernest, C. Steven; Shcherbinin, Sergey et al. (2023): Donanemab Population Pharmacokinetics, Amyloid Plaque Reduction, and Safety in

Participants with Alzheimer's Disease. In *Clinical pharmacology and therapeutics* 113 (6), pp. 1258–1267. DOI: 10.1002/cpt.2875.

Hanagasi, Hasmet A.; Tufekcioglu, Zeynep; Emre, Murat (2017): Dementia in Parkinson's disease. In *Journal of the neurological sciences* 374, pp. 26–31. DOI: 10.1016/j.jns.2017.01.012.

Harrell, Frank E. (1982): Evaluating the Yield of Medical Tests. In *JAMA* 247 (18), p. 2543. DOI: 10.1001/jama.1982.03320430047030.

Harris, M. I.; Klein, R.; Welborn, T. A.; Knuiman, M. W. (1992): Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. In *Diabetes care* 15 (7), pp. 815–819. DOI: 10.2337/diacare.15.7.815.

Hayes-Larson, Eleanor; Ikesu, Ryo; Fong, Joseph; Mobley, Taylor M.; Gee, Gilbert C.; Brookmeyer, Ron et al. (2023): Association of Education With Dementia Incidence Stratified by Ethnicity and Nativity in a Cohort of Older Asian American Individuals. In *JAMA network open* 6 (3), e231661. DOI: 10.1001/jamanetworkopen.2023.1661.

Hayward, Mark D.; Farina, Mateo P.; Zhang, Yuan S.; Kim, Jung Ki; Crimmins, Eileen M. (2021): The Importance of Improving Educational Attainment for Dementia Prevalence Trends From 2000 to 2014, Among Older Non-Hispanic Black and White Americans. In *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 76 (9), pp. 1870–1879. DOI: 10.1093/geronb/gbab015.

Hendriks, Stevie; Peetoom, Kirsten; Bakker, Christian; van der Flier, Wiesje M.; Papma, Janne M.; Koopmans, Raymond et al. (2021): Global Prevalence of Young-Onset Dementia: A Systematic Review and Meta-analysis. In *JAMA neurology* 78 (9), pp. 1080–1090. DOI: 10.1001/jamaneurol.2021.2161.

Hennessy, Sean (2006): Use of health care databases in pharmacoepidemiology. In *Basic & clinical pharmacology & toxicology* 98 (3), pp. 311–313. DOI: 10.1111/j.1742-7843.2006.pto_368.x.

Herd, Pamela; Sicinski, Kamil (2022): Using sibling models to unpack the relationship between education and cognitive functioning in later life. In *SSM - Population Health* 17, p. 100960. DOI: 10.1016/j.ssmph.2021.100960.

Hernán, Miguel A.; Sauer, Brian C.; Hernández-Díaz, Sonia; Platt, Robert; Shrier, Ian (2016): Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. In *Journal of clinical epidemiology* 79, pp. 70–75. DOI: 10.1016/j.jclinepi.2016.04.014.

Heser, K.; Fink, A.; Reinke, C.; Wagner, M.; Doblhammer, G. (2020): The temporal association between incident late-life depression and incident dementia. In *Acta psychiatrica Scandinavica* 142 (5), pp. 402–412. DOI: 10.1111/acps.13220.

HHS (Ed.) (2023): National Plan to Address Alzheimer's Disease: 2023 Update. U.S. Department of Health and Human Services. Available online at https://aspe.hhs.gov/collaborations-committees-advisory-groups/napa/napa-documents/napa-national-plan, checked on 11/22/2024.

Hoffmann, Falk (2009): Review on use of German health insurance medication claims data for epidemiological research. In *Pharmacoepidemiology and drug safety* 18 (5), pp. 349–356. DOI: 10.1002/pds.1721.

Høilund-Carlsen, Poul F.; Alavi, Abass; Barrio, Jorge R.; Castellani, Rudolph J.; Costa, Tommaso; Herrup, Karl et al. (2024): Donanemab, another anti-Alzheimer's drug with risk and uncertain benefit. In *Ageing research reviews* 99, p. 102348. DOI: 10.1016/j.arr.2024.102348.

Hopkins, Rachel; Shaver, Kristi; Weinstock, Ruth S. (2016): Management of Adults With Diabetes and Cognitive Problems. In *Diabetes spectrum : a publication of the American Diabetes Association* 29 (4), pp. 224–237. DOI: 10.2337/ds16-0035.

Hou, Xiao-He; Feng, Lei; Zhang, Can; Cao, Xi-Peng; Tan, Lan; Yu, Jin-Tai (2019): Models for predicting risk of dementia: a systematic review. In *Journal of neurology, neurosurgery, and psychiatry* 90 (4), pp. 373–379. DOI: 10.1136/jnnp-2018-318212.

Hsiao, Julia J.; Teng, Edmond (2013): Depressive Symptoms in Clinical and Incipient Alzheimer's Disease. In *Neurodegenerative Disease Management* 3 (2), pp. 147–155. DOI: 10.2217/nmt.13.10.

Hughes, Diarmaid; Judge, Conor; Murphy, Robert; Loughlin, Elaine; Costello, Maria; Whiteley, William et al. (2020): Association of Blood Pressure Lowering With Incident Dementia or Cognitive Impairment: A Systematic Review and Meta-analysis. In *JAMA* 323 (19), pp. 1934–1944. DOI: 10.1001/jama.2020.4249.

Husain-Krautter, Sehba; Ellison, James M. (2021): Late Life Depression: The Essentials and the Essential Distinctions. In *Focus (American Psychiatric Publishing)* 19 (3), pp. 282–293. DOI: 10.1176/appi.focus.20210006.

Huuha, Aleksi M.; Norevik, Cecilie S.; Moreira, José Bianco N.; Kobro-Flatmoen, Asgeir; Scrimgeour, Nathan; Kivipelto, Miia et al. (2022): Can exercise training teach us how to treat Alzheimer's disease? In *Ageing research reviews* 75, p. 101559. DOI: 10.1016/j.arr.2022.101559.

Hyun, Jinshil; Hall, Charles B.; Katz, Mindy J.; Derby, Carol A.; Lipnicki, Darren M.; Crawford, John D. et al. (2022): Education, Occupational Complexity, and Incident Dementia: A COSMIC Collaborative Cohort Study. In *JAD* 85 (1), pp. 179–196. DOI: 10.3233/JAD-210627.

Imai, Kosuke; Keele, Luke; Tingley, Dustin (2010a): A general approach to causal mediation analysis. In *Psychological methods* 15 (4), pp. 309–334. DOI: 10.1037/a0020761.

Imai, Kosuke; Keele, Luke; Yamamoto, Teppei (2010b): Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. In *Statist. Sci.* 25 (1). DOI: 10.1214/10-STS321.

Jacobs, E.; Rathmann, W.; Tönnies, T.; Arendt, D.; Marchowez, M.; Veith, L. et al. (2020): Age at diagnosis of Type 2 diabetes in Germany: a nationwide analysis based on claims data from 69 million people. In *Diabetic medicine : a journal of the British Diabetic Association* 37 (10), pp. 1723–1727. DOI: 10.1111/dme.14100.

Jain, Siddharth; Rosenbaum, Paul R.; Reiter, Joseph G.; Hoffman, Geoffrey; Small, Dylan S.; Ha, Jinkyung et al. (2020): Using Medicare claims in identifying Alzheimer's disease and related dementias. In *Alzheimer's & Dementia*. DOI: 10.1002/alz.12199.

Jayaraj, Richard L.; Rodriguez, Eric A.; Wang, Yi; Block, Michelle L. (2017): Outdoor Ambient Air Pollution and Neurodegenerative Diseases: the Neuroinflammation Hypothesis. In *Current environmental health reports* 4 (2), pp. 166–179. DOI: 10.1007/s40572-017-0142-3.

Jin, Jiayin; Sommerlad, Andrew; Mukadam, Naaheed (2024): Association between adult education, brain volume and dementia risk: longitudinal cohort study of UK Biobank participants. In *GeroScience*. DOI: 10.1007/s11357-024-01285-y.

Kandimalla, Ramesh; Thirumala, Vani; Reddy, P. Hemachandra (2017): Is Alzheimer's disease a Type 3 Diabetes? A critical appraisal. In *Biochimica et biophysica acta. Molecular basis of disease* 1863 (5), pp. 1078–1089. DOI: 10.1016/j.bbadis.2016.08.018.

Karlsson, Ida K.; Bennet, Anna M.; Ploner, Alexander; Andersson, Therese M-L; Reynolds, Chandra A.; Gatz, Margaret; Pedersen, Nancy L. (2015): Apolipoprotein E ε4 genotype and the temporal relationship between depression and dementia. In *Neurobiology of aging* 36 (4), pp. 1751–1756. DOI: 10.1016/j.neurobiolaging.2015.01.008.

Karr, Justin E.; Graham, Raquel B.; Hofer, Scott M.; Muniz-Terrera, Graciela (2018): When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death. In *Psychology and Aging* 33 (2), pp. 195–218. DOI: 10.1037/pag0000236.

Katz, Douglas I.; Bernick, Charles; Dodick, David W.; Mez, Jesse; Mariani, Megan L.; Adler, Charles H. et al. (2021): National Institute of Neurological Disorders and Stroke Consensus Diagnostic Criteria for Traumatic Encephalopathy Syndrome. In *Neurology* 96 (18), pp. 848–863. DOI: 10.1212/WNL.0000000000011850.

Kelly, Michelle E.; Duff, Hollie; Kelly, Sara; McHugh Power, Joanna E.; Brennan, Sabina; Lawlor, Brian A.; Loughrey, David G. (2017): The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. In *Systematic reviews* 6 (1), p. 259. DOI: 10.1186/s13643-017-0632-2.

Khan, Moien Abdul Basith; Hashim, Muhammad Jawad; King, Jeffrey Kwan; Govender, Romona Devi; Mustafa, Halla; Al Kaabi, Juma (2020): Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends. In *JEGH* 10 (1), p. 107. DOI: 10.2991/jegh.k.191028.001.

Kiely, Kim M. (2014): Cognitive Function. In Alex C. Michalos (Ed.): Encyclopedia of quality of life and well-being research. Dordrecht: Springer, pp. 974–978.

Kirch, Wilhelm (Ed.) (2008): Encyclopedia of public health. New York, NY, Dordrecht: Springer; Springer Netherlands Online-Anbieter.

Kivimäki, Mika; Singh-Manoux, Archana; Batty, G. David; Sabia, Séverine; Sommerlad, Andrew; Floud, Sarah et al. (2020): Association of Alcohol-Induced Loss of Consciousness and Overall Alcohol Consumption With Risk for Dementia. In *JAMA network open* 3 (9), e2016084. DOI: 10.1001/jamanetworkopen.2020.16084.

Kivimäki, Mika; Walker, Keenan A.; Pentti, Jaana; Nyberg, Solja T.; Mars, Nina; Vahtera, Jussi et al. (2021): Cognitive stimulation in the workplace, plasma proteins, and risk of dementia: three analyses of population cohort studies. In *BMJ (Clinical research ed.)* 374, n1804. DOI: 10.1136/bmj.n1804.

Kjeldsen, E. W.; Thomassen, J. Q.; Rasmussen, K. L.; Nordestgaard, B. G.; Tybjærg-Hansen, A.; Frikke-Schmidt, R. (2022): Adherence to dietary guidelines and risk of dementia: a prospective cohort study of 94 184 individuals. In *Epidemiology and psychiatric sciences* 31, e71. DOI: 10.1017/S2045796022000567.

Klein, Eric G.; Schroeder, Krista; Wessels, Alette M.; Phipps, Adam; Japha, Maureen; Schilling, Traci; Zimmer, Jennifer A. (2024): How donanemab data address the coverage with evidence development questions. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 20 (4), pp. 3127–3140. DOI: 10.1002/alz.13700.

Kowall, Bernd; Rathmann, Wolfgang (2023): Combined Effects of Diabetes and Education on Decline of Cognitive Performance in the Older Population: The Survey of Health, Ageing, and Retirement in Europe. In *Gerontology* 69 (2), pp. 172–180. DOI: 10.1159/000524571.

Kreis, Kristine; Neubauer, Sarah; Klora, Mike; Lange, Ansgar; Zeidler, Jan (2016): Status and perspectives of claims data analyses in Germany-A systematic review. In *Health policy (Amsterdam, Netherlands)* 120 (2), pp. 213–226. DOI: 10.1016/j.healthpol.2016.01.007.

Kuh, D.; Ben-Shlomo, Y.; Lynch, J.; Hallqvist, J.; Power, C. (2003): Life course epidemiology. In *Journal of epidemiology and community health* 57 (10), pp. 778–783. DOI: 10.1136/jech.57.10.778.

Kulzer, B.; Albus, C.; Herpertz, S.; Kruse, J.; Lange, K.; Lederbogen, F.; Petrak, F. (2013): Psychosoziales und Diabetes (Teil 1). In *Diabetologie und Stoffwechsel* 8 (03), pp. 198–242. DOI: 10.1055/s-0033-1335785.

Kuźma, Elżbieta; Littlejohns, Thomas J.; Khawaja, Anthony P.; Llewellyn, David J.; Ukoumunne, Obioha C.; Thiem, Ulrich (2021): Visual Impairment, Eye Diseases, and Dementia Risk: A Systematic Review and Meta-Analysis. In *JAD* 83 (3), pp. 1073–1087. DOI: 10.3233/JAD-210250.

La Monte, Suzanne M. de; Tong, Ming; Wands, Jack R. (2018): The 20-Year Voyage Aboard the Journal of Alzheimer's Disease: Docking at 'Type 3 Diabetes', Environmental/Exposure Factors, Pathogenic Mechanisms, and Potential Treatments. In *JAD* 62 (3), pp. 1381–1390. DOI: 10.3233/JAD-170829.

Li, Fu-Rong; Yang, Hai-Lian; Zhou, Rui; Zheng, Jia-Zhen; Chen, Guo-Chong; Wu, Xiao-Xiang et al. (2021): Influence of Diabetes Duration and Glycemic Control on Dementia: A Cohort Study. In *The journals of gerontology. Series A, Biological sciences and medical sciences* 76 (11), pp. 2062–2070. DOI: 10.1093/gerona/glab221.

Lifelines (2024): Mandatory statement. For findability and visibility of Lifelines. Edited by Lifelines. Available online at https://www.lifelines-biobank.com/researchers/frequently-asked-questions/mandatory-content/findability-and-visibility, checked on 10/18/2024.

Lim, Yen Ying; Jaeger, Judith; Harrington, Karra; Ashwood, Tim; Ellis, Kathryn A.; Stöffler, Albrecht et al. (2013): Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian Imaging, Biomarkers, and Lifestyle-rate of change substudy (AIBL-ROCS). In *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists* 28 (4), pp. 320–330. DOI: 10.1093/arclin/act021.

Liu, Chelsea; Murchland, Audrey R.; VanderWeele, Tyler J.; Blacker, Deborah (2022): Eliminating racial disparities in dementia risk by equalizing education quality: A sensitivity analysis. In *Social science & medicine (1982)* 312, p. 115347. DOI: 10.1016/j.socscimed.2022.115347.

Liu, Yi-Hsuan; Gao, Xiang; Na, Muzi; Kris-Etherton, Penny M.; Mitchell, Diane C.; Jensen, Gordon L. (2020): Dietary Pattern, Diet Quality, and Dementia: A Systematic Review and Meta-Analysis of Prospective Cohort Studies. In *JAD* 78 (1), pp. 151–168. DOI: 10.3233/JAD-200499.

Liu, Yulu; Lu, Guangyu; Liu, Lin; He, Yuhang; Gong, Weijuan (2024): Cognitive reserve over the life course and risk of dementia: a systematic review and meta-analysis. In *Frontiers in aging neuroscience* 16, p. 1358992. DOI: 10.3389/fnagi.2024.1358992.

Livingston, Gill; Huntley, Jonathan; Liu, Kathy Y.; Costafreda, Sergi G.; Selbæk, Geir; Alladi, Suvarna et al. (2024): Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. In *The Lancet*. DOI: 10.1016/S0140-6736(24)01296-0.

Livingston, Gill; Huntley, Jonathan; Sommerlad, Andrew; Ames, David; Ballard, Clive; Banerjee, Sube et al. (2020): Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. In *Lancet (London, England)* 396 (10248), pp. 413–446. DOI: 10.1016/S0140-6736(20)30367-6.

Lövdén, Martin; Fratiglioni, Laura; Glymour, M. Maria; Lindenberger, Ulman; Tucker-Drob, Elliot M. (2020): Education and Cognitive Functioning Across the Life Span. In *Psychological science in the public interest : a journal of the American Psychological Society* 21 (1), pp. 6–41. DOI: 10.1177/1529100620920576.

Loy, Clement T.; Schofield, Peter R.; Turner, Anne M.; Kwok, John B. J. (2014): Genetics of dementia. In *Lancet (London, England)* 383 (9919), pp. 828–840. DOI: 10.1016/S0140-6736(13)60630-3.

Lynch, J. W.; Kaplan, G. A.; Salonen, J. T. (1997): Why do poor people behave poorly? Variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic lifecourse. In *Social science & medicine (1982)* 44 (6), pp. 809–819. DOI: 10.1016/S0277-9536(96)00191-8.

Mahinrad, Simin; Bennett, David A.; Sorond, Farzaneh A.; Gorelick, Philip B. (2023): Blood pressure variability, dementia, and role of antihypertensive medications in older adults. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 19 (7), pp. 2966–2974. DOI: 10.1002/alz.12935.

Marawi, Tulip; Ainsworth, Nicholas J.; Zhukovsky, Peter; Rashidi-Ranjbar, Neda; Rajji, Tarek K.; Tartaglia, Maria Carmela et al. (2023): Brain-cognition relationships in late-life depression: a systematic review of structural magnetic resonance imaging studies. In *Translational psychiatry* 13 (1), p. 284. DOI: 10.1038/s41398-023-02584-2.

Maruff, Paul; Lim, Yen Ying; Darby, David; Ellis, Kathryn A.; Pietrzak, Robert H.; Snyder, Peter J. et al. (2013): Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer's disease. In *BMC psychology* 1 (1), p. 30. DOI: 10.1186/2050-7283-1-30.

McIntosh, Elissa C.; Nation, Daniel A. (2019): Importance of Treatment Status in Links Between Type 2 Diabetes and Alzheimer's Disease. In *Diabetes care* 42 (5), pp. 972–979. DOI: 10.2337/dc18-1399.

McKhann, Guy M.; Knopman, David S.; Chertkow, Howard; Hyman, Bradley T.; Jack, Clifford R.; Kawas, Claudia H. et al. (2011): The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 7 (3), pp. 263–269. DOI: 10.1016/j.jalz.2011.03.005.

Meng, Xiangfei; D'Arcy, Carl (2012): Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. In *PloS one* 7 (6), e38268. DOI: 10.1371/journal.pone.0038268.

Miller, M. E.; Langefeld, C. D.; Tierney, W. M.; Hui, S. L.; McDonald, C. J. (1993): Validation of probabilistic predictions. In *Medical decision making : an international journal of the Society for Medical Decision Making* 13 (1), pp. 49–58. DOI: 10.1177/0272989X9301300107.

Mirza, Saira Saeed; Bruijn, Renée F. A. G. de; Direk, Nese; Hofman, Albert; Koudstaal, Peter J.; Ikram, M. Arfan; Tiemeier, Henning (2014): Depressive symptoms predict incident dementia during short- but not long-term follow-up period. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 10 (5 Suppl), S323-S329.e1. DOI: 10.1016/j.jalz.2013.10.006.

Mirza, Saira Saeed; Wolters, Frank J.; Swanson, Sonja A.; Koudstaal, Peter J.; Hofman, Albert; Tiemeier, Henning; Ikram, M. Arfan (2016): 10-year trajectories of depressive symptoms and risk of dementia: a population-based study. In *The lancet. Psychiatry* 3 (7), pp. 628–635. DOI: 10.1016/S2215-0366(16)00097-3.

Monette, Michelle C. E.; Baird, Anne; Jackson, Dennis L. (2014): A meta-analysis of cognitive functioning in nondemented adults with type 2 diabetes mellitus. In *Canadian journal of diabetes* 38 (6), pp. 401–408. DOI: 10.1016/j.jcjd.2014.01.014.

Müters, S. (2013): Kartografische Darstellung für den Bericht „Daten und Fakten: Ergebnisse der Studie Gesundheit in Deutschland aktuell 2010“. In *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))* 75 (04). DOI: 10.1055/s-0033-1337460.

Nakahori, Nobue; Sekine, Michikazu; Yamada, Masaaki; Tatsuse, Takashi; Kido, Hideki; Suzuki, Michio (2018): A pathway from low socioeconomic status to dementia in Japan: results from the Toyama dementia survey. In *BMC geriatrics* 18 (1), p. 102. DOI: 10.1186/s12877-018-0791-6.

Nelson, Monica E.; Jester, Dylan J.; Petkus, Andrew J.; Andel, Ross (2021): Cognitive Reserve, Alzheimer's Neuropathology, and Risk of Dementia: A Systematic Review and Meta-Analysis. In *Neuropsychology review* 31 (2), pp. 233–250. DOI: 10.1007/s11065-021-09478-4.

Nelson, Peter T.; Head, Elizabeth; Schmitt, Frederick A.; Davis, Paulina R.; Neltner, Janna H.; Jicha, Gregory A. et al. (2011): Alzheimer's disease is not "brain aging": neuropathological, genetic, and epidemiological human studies. In *Acta neuropathologica* 121 (5), pp. 571–587. DOI: 10.1007/s00401-011-0826-y.

Nerius, M.; Fink, A.; Doblhammer, G. (2017): Parkinson's disease in Germany: prevalence and incidence based on health claims data. In *Acta neurologica Scandinavica* 136 (5), pp. 386–392. DOI: 10.1111/ane.12694.

Nerius, Michael; Ziegler, Uta; Doblhammer, Gabriele; Fink, Anne (2020): Trends in der Prävalenz von Demenz und Parkinson – Eine Analyse auf Basis aller gesetzlich versicherten Personen im Alter 65+in Deutschland zwischen 2009 und 2012. In *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))* 82 (10), pp. 761–769. DOI: 10.1055/a-0829-6494.

Nichols et al., Emma (2022): Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. In *The Lancet. Public health* 7 (2), e105-e125. DOI: 10.1016/S2468-2667(21)00249-8.

Nogueira, Joana; Gerardo, Bianca; Santana, Isabel; Simões, Mário R.; Freitas, Sandra (2022): The Assessment of Cognitive Reserve: A Systematic Review of the Most Used Quantitative Measurement Methods of Cognitive Reserve for Aging. In *Frontiers in psychology* 13, p. 847186. DOI: 10.3389/fpsyg.2022.847186.

Nori, Vijay S.; Hane, Christopher A.; Martin, David C.; Kravetz, Alexander D.; Sanghavi, Darshak M. (2019): Identifying incident dementia by applying machine learning to a very large administrative claims dataset. In *PloS one* 14 (7), e0203246. DOI: 10.1371/journal.pone.0203246.

OECD (2023): Measuring Population Mental Health. Paris: OECD Publishing.

Offord, D. R.; Kraemer, H. C. (2000): Risk factors and prevention. In *Evidence-Based Mental Health* 3 (3), pp. 70–71. DOI: 10.1136/ebmh.3.3.70.

Ott, A.; Stolk, R. P.; van Harskamp, F.; Pols, H. A.; Hofman, A.; Breteler, M. M. (1999): Diabetes mellitus and the risk of dementia: The Rotterdam Study. In *Neurology* 53 (9), pp. 1937–1942. DOI: 10.1212/wnl.53.9.1937.

Paik, Ji-Sun; Ha, Minji; Jung, Youn Hea; Kim, Gee-Hyun; Han, Kyung-Do; Kim, Hyun-Seung et al. (2020): Low vision and the risk of dementia: a nationwide population-based cohort study. In *Scientific reports* 10 (1), p. 9109. DOI: 10.1038/s41598-020-66002-z.

Palta, Priya; Schneider, Andrea L. C.; Biessels, Geert Jan; Touradji, Pegah; Hill-Briggs, Felicia (2014): Magnitude of cognitive dysfunction in adults with type 2 diabetes: a meta-analysis of six cognitive domains and the most frequently reported neuropsychological tests within domains. In *Journal of the International Neuropsychological Society : JINS* 20 (3), pp. 278–291. DOI: 10.1017/S1355617713001483.

Pandya, Seema Y.; Clem, Matthew A.; Silva, Lynette M.; Woon, Fu L. (2016): Does mild cognitive impairment always lead to dementia? A review. In *Journal of the neurological sciences* 369, pp. 57–62. DOI: 10.1016/j.jns.2016.07.055.

Panico, Francesco; Sagliano, Laura; Magliacano, Alfonso; Santangelo, Gabriella; Trojano, Luigi (2023): The relationship between cognitive reserve and cognition in healthy adults: a systematic review. In *Curr Psychol* 42 (28), pp. 24751–24763. DOI: 10.1007/s12144-022-03523-y.

Pappalettera, Chiara; Carrarini, Claudia; Miraglia, Francesca; Vecchio, Fabrizio; Rossini, Paolo M. (2024): Cognitive resilience/reserve: Myth or reality? A review of definitions and measurement methods. In *Alzheimer's & Dementia* 20 (5), pp. 3567–3586. DOI: 10.1002/alz.13744.

Parra, Kimberly L.; Alexander, Gene E.; Raichlen, David A.; Klimentidis, Yann C.; Furlong, Melissa A. (2022): Exposure to air pollution and risk of incident dementia in the UK Biobank. In *Environmental research* 209, p. 112895. DOI: 10.1016/j.envres.2022.112895.

Peña-González, Paola; Mondragón-Maya, Alejandra; Silva-Pereyra, Juan; Roa-Rojas, Paloma (2020): Cognitive Reserve and Executive Functions in Adults with Type 2 Diabetes. In *Journal of Diabetes Research* 2020, pp. 1–7. DOI: 10.1155/2020/7941543.

Pérès, Karine; Brayne, Carol; Matharan, Fanny; Grasset, Leslie; Helmer, Catherine; Letenneur, Luc et al. (2017): Trends in Prevalence of Dementia in French Farmers from Two Epidemiological Cohorts. In *Journal of the American Geriatrics Society* 65 (2), pp. 415–420. DOI: 10.1111/jgs.14575.

Peters, Ruth; Mudway, Ian; Booth, Andrew; Peters, Jean; Anstey, Kaarin J. (2021): Putting Fine Particulate Matter and Dementia in the Wider Context of Noncommunicable Disease: Where are We Now and What Should We Do Next: A Systematic Review. In *Neuroepidemiology* 55 (4), pp. 253–265. DOI: 10.1159/000515394.

Peters, Ruth; Poulter, Ruth; Warner, James; Beckett, Nigel; Burch, Lisa; Bulpitt, Chris (2008): Smoking, dementia and cognitive decline in the elderly, a systematic review. In *BMC geriatrics* 8, p. 36. DOI: 10.1186/1471-2318-8-36.

Pettigrew, Corinne; Soldan, Anja (2019): Defining Cognitive Reserve and Implications for Cognitive Aging. In *Current neurology and neuroscience reports* 19 (1), p. 1. DOI: 10.1007/s11910-019-0917-z.

Powell, Danielle S.; Brenowitz, Willa D.; Yaffe, Kristine; Armstrong, Nicole M.; Reed, Nicholas S.; Lin, Frank R. et al. (2022): Examining the Combined Estimated Effects of Hearing Loss and Depressive Symptoms on Risk of Cognitive Decline and Incident Dementia. In *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 77 (5), pp. 839–849. DOI: 10.1093/geronb/gbab194.

Puttanna, Amar; Padinjakara, Noushad Kulambil (2017): Management of diabetes and dementia. In *Br J Diabetes* 17 (3), p. 93. DOI: 10.15277/bjd.2017.139.

Qiao, Yue; Gu, Jian; Yu, Miao; Chi, Yuewei; Ma, Ying (2024): Comparative Efficacy and Safety of Monoclonal Antibodies for Cognitive Decline in Patients with Alzheimer's Disease: A Systematic Review and Network Meta-Analysis. In *CNS drugs* 38 (3), pp. 169–192. DOI: 10.1007/s40263-024-01067-2.

Qu, Yi; Hu, He-Ying; Ou, Ya-Nan; Shen, Xue-Ning; Xu, Wei; Wang, Zuo-Teng et al. (2020): Association of body mass index with risk of cognitive impairment and dementia: A systematic review and meta-analysis of prospective studies. In *Neuroscience and biobehavioral reviews* 115, pp. 189–198. DOI: 10.1016/j.neubiorev.2020.05.012.

Raggi, Martina; Dugravot, Aline; Valeri, Linda; Machado-Fragua, Marcos D.; Dumurgier, Julien; Kivimaki, Mika et al. (2022): Contribution of smoking towards the association between socioeconomic position and dementia: 32-year follow-up of the Whitehall II prospective cohort study. In *The Lancet regional health. Europe* 23, p. 100516. DOI: 10.1016/j.lanepe.2022.100516.

Rakuša, Elena; Fink, Anne; Tamgüney, Gültekin; Heneka, Michael T.; Doblhammer, Gabriele (2023): Sporadic Use of Antibiotics in Older Adults and the Risk of Dementia: A Nested Case-Control Study Based on German Health Claims Data. In *JAD* 93 (4), pp. 1329–1339. DOI: 10.3233/JAD-221153.

Raz, Limor; Knoefel, Janice; Bhaskar, Kiran (2016): The neuropathology and cerebrovascular mechanisms of dementia. In *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* 36 (1), pp. 172–186. DOI: 10.1038/jcbfm.2015.164.

Rehm, Jürgen; Hasan, Omer S. M.; Black, Sandra E.; Shield, Kevin D.; Schwarzinger, Michaël (2019): Alcohol use and dementia: a systematic scoping review. In *Alzheimer's research & therapy* 11 (1), p. 1. DOI: 10.1186/s13195-018-0453-0.

Reinke, Constantin (2024): The effect of diabetes in the multifaceted relationship between education and cognitive function. In *BMC Public Health* 24 (1). DOI: 10.1186/s12889-024-20156-x.

Reinke, Constantin; Buchmann, Nikolaus; Fink, Anne; Tegeler, Christina; Demuth, Ilja; Doblhammer, Gabriele (2022): Diabetes duration and the risk of dementia: a cohort study based on German health claims data. In *Age and ageing* 51 (1). DOI: 10.1093/ageing/afab231.

Reinke, Constantin; Doblhammer, Gabriele; Schmid, Matthias; Welchowski, Thomas (2023): Dementia risk predictions from German claims data using methods of machine learning. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 19 (2), pp. 477–486. DOI: 10.1002/alz.12663.

Reynolds, Charles F.; Lenze, Eric; Mulsant, Benoit H. (2019): Assessment and treatment of major depression in older adults. In *Handbook of clinical neurology* 167, pp. 429–435. DOI: 10.1016/B978-0-12-804766-8.00023-6.

Rizzuto, Debora; Feldman, Adina L.; Karlsson, Ida K.; Dahl Aslan, Anna K.; Gatz, Margaret; Pedersen, Nancy L. (2018): Detection of Dementia Cases in Two Swedish Health Registers: A Validation Study. In *JAD* 61 (4), pp. 1301–1310. DOI: 10.3233/JAD-170572.

Rolandi, Elena; Zaccaria, Daniele; Vaccaro, Roberta; Abbondanza, Simona; Pettinato, Laura; Davin, Annalisa; Guaita, Antonio (2020): Estimating the potential for dementia prevention through modifiable risk factors elimination in the real-world setting: a population-based study. In *Alzheimer's research & therapy* 12 (1), p. 94. DOI: 10.1186/s13195-020-00661-y.

Rorbach-Dolata, Anna; Piwowar, Agnieszka (2019): Neurometabolic Evidence Supporting the Hypothesis of Increased Incidence of Type 3 Diabetes Mellitus in the 21st Century. In *BioMed research international* 2019, p. 1435276. DOI: 10.1155/2019/1435276.

Sachdev, Perminder S.; Blacker, Deborah; Blazer, Dan G.; Ganguli, Mary; Jeste, Dilip V.; Paulsen, Jane S.; Petersen, Ronald C. (2014): Classifying neurocognitive disorders: the DSM-5 approach. In *Nature reviews. Neurology* 10 (11), pp. 634–642. DOI: 10.1038/nrneurol.2014.181.

Safaei, Mahmood; Sundararajan, Elankovan A.; Driss, Maha; Boulila, Wadii; Shapi'i, Azrulhizam (2021): A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. In *Computers in biology and medicine* 136, p. 104754. DOI: 10.1016/j.compbiomed.2021.104754.

Saiepour, N.; Najman, J. M.; Ware, R.; Baker, P.; Clavarino, A. M.; Williams, G. M. (2019): Does attrition affect estimates of association: A longitudinal study. In *Journal of psychiatric research* 110, pp. 127–142. DOI: 10.1016/j.jpsychires.2018.12.022.

Sapsford, Timothy P.; Johnson, Susannah R.; Headrick, John P.; Branjerdporn, Grace; Adhikary, Sam; Sarfaraz, Muhammad; Stapelberg, Nicolas J. C. (2022): Forgetful, sad and old: Do vascular cognitive impairment and depression share a common pre-disease network and how is it impacted by ageing? In *Journal of psychiatric research* 156, pp. 611–627. DOI: 10.1016/j.jpsychires.2022.10.071.

Schmid, Matthias; Tutz, Gerhard; Welchowski, Thomas (2018): Discrimination measures for discrete time-to-event predictions. In *Econometrics and Statistics* 7, pp. 153–164. DOI: 10.1016/j.ecosta.2017.03.008.

Schneeweiss, Sebastian; Avorn, Jerry (2005): A review of uses of health care utilization databases for epidemiologic research on therapeutics. In *Journal of clinical epidemiology* 58 (4), pp. 323–337. DOI: 10.1016/j.jclinepi.2004.10.012.

Scholtens, Salome; Smidt, Nynke; Swertz, Morris A.; Bakker, Stephan J. L.; Dotinga, Aafje; Vonk, Judith M. et al. (2015): Cohort Profile: LifeLines, a three-generation cohort study and biobank. In *International Journal of Epidemiology* 44 (4), pp. 1172–1180. DOI: 10.1093/ije/dyu229.

Schubert, I.; Ihle, P.; Köster, I. (2010): Interne Validierung von Diagnosen in GKV-Routinedaten: Konzeption mit Beispielen und Falldefinition. In *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))* 72 (6), pp. 316–322. DOI: 10.1055/s-0030-1249688.

Seblova, D.; Berggren, R.; Lövdén, M. (2020): Education and age-related decline in cognitive performance: Systematic review and meta-analysis of longitudinal cohort studies. In *Ageing research reviews* 58, p. 101005. DOI: 10.1016/j.arr.2019.101005.

Seblova, Dominika; Fischer, Martin; Fors, Stefan; Johnell, Kristina; Karlsson, Martin; Nilsson, Therese et al. (2021): Does Prolonged Education Causally Affect Dementia Risk When Adult Socioeconomic Status Is Not Altered? A Swedish Natural Experiment in 1.3 Million Individuals. In *American journal of epidemiology* 190 (5), pp. 817–826. DOI: 10.1093/aje/kwaa255.

Shang, Xianwen; Zhu, Zhuoting; Huang, Yu; Zhang, Xueli; Wang, Wei; Shi, Danli et al. (2023): Associations of ophthalmic and systemic conditions with incident dementia in the UK Biobank. In *The British journal of ophthalmology* 107 (2), pp. 275–282. DOI: 10.1136/bjophthalmol-2021-319508.

Shang, Xianwen; Zhu, Zhuoting; Wang, Wei; Ha, Jason; He, Mingguang (2021): The Association between Vision Impairment and Incidence of Dementia and Cognitive Impairment. In *Ophthalmology* 128 (8), pp. 1135–1149. DOI: 10.1016/j.ophtha.2020.12.029.

Shen, Chun; Rolls, Edmund T.; Cheng, Wei; Kang, Jujiao; Dong, Guiying; Xie, Chao et al. (2022): Associations of Social Isolation and Loneliness With Later Dementia. In *Neurology* 99 (2), e164-e175. DOI: 10.1212/WNL.0000000000200583.

Shrivastava, Saurabh Rambiharilal; Shrivastava, Prateek Saurabh; Ramasamy, Jegadeesh (2013): Role of self-care in management of diabetes mellitus. In *Journal of diabetes and metabolic disorders* 12 (1), p. 14. DOI: 10.1186/2251-6581-12-14.

Shukla, Aishwarya; Harper, Michael; Pedersen, Emily; Goman, Adele; Suen, Jonathan J.; Price, Carrie et al. (2020): Hearing Loss, Loneliness, and Social Isolation: A Systematic Review. In *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 162 (5), pp. 622–633. DOI: 10.1177/0194599820910377.

Sijtsma, Anna; Rienks, Johanna; van der Harst, Pim; Navis, Gerjan; Rosmalen, Judith G. M.; Dotinga, Aafje (2022): Cohort Profile Update: Lifelines, a three-generation cohort study and biobank. In *International Journal of Epidemiology* 51 (5), e295-e302. DOI: 10.1093/ije/dyab257.

Sims, John R.; Zimmer, Jennifer A.; Evans, Cynthia D.; Lu, Ming; Ardayfio, Paul; Sparks, JonDavid et al. (2023): Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. In *JAMA* 330 (6), pp. 512–527. DOI: 10.1001/jama.2023.13239.

Sinclair, Lindsey I.; Mohr, Asher; Morisaki, Mizuki; Edmondson, Martin; Chan, Selina; Bone-Connaughton, A. et al. (2023): Is later-life depression a risk factor for Alzheimer's disease or a prodromal symptom: a study using post-mortem human brain tissue? In *Alzheimer's research & therapy* 15 (1), p. 153. DOI: 10.1186/s13195-023-01299-2.

Singh-Manoux, Archana; Dugravot, Aline; Fournier, Agnes; Abell, Jessica; Ebmeier, Klaus; Kivimäki, Mika; Sabia, Séverine (2017): Trajectories of Depressive Symptoms Before Diagnosis of Dementia: A 28-Year Follow-up Study. In *JAMA psychiatry* 74 (7), pp. 712–718. DOI: 10.1001/jamapsychiatry.2017.0660.

Soh, Yenee; Whitmer, Rachel A.; Mayeda, Elizabeth Rose; Glymour, M. Maria; Eng, Chloe W.; Peterson, Rachel L. et al. (2024): Timing and level of educational attainment and late-life cognition in the KHANDLE study. In *Alzheimer's & Dementia* 20 (1), pp. 593–600. DOI: 10.1002/alz.13475.

Song, Tanya; Wang, Yunfei; Silverglate, Bret David; Grossberg, George T. (2024): Pharmacokinetic evaluation of donanemab for the treatment of Alzheimer's. In *Expert opinion on drug metabolism & toxicology* 20 (6), pp. 411–417. DOI: 10.1080/17425255.2024.2357637.

Stafford, Jean; Chung, Wing Tung; Sommerlad, Andrew; Kirkbride, James B.; Howard, Robert (2022): Psychiatric disorders and risk of subsequent dementia: Systematic review and meta-analysis of longitudinal studies. In *International journal of geriatric psychiatry* 37 (5). DOI: 10.1002/gps.5711.

Steen, Eric; Terry, Benjamin M.; Rivera, Enrique J.; Cannon, Jennifer L.; Neely, Thomas R.; Tavares, Rose et al. (2005): Impaired insulin and insulin-like growth factor expression and signaling mechanisms in Alzheimer's disease--is this type 3 diabetes? In *JAD* 7 (1), pp. 63–80. DOI: 10.3233/jad-2005-7107.

Stern, Yaakov (2009): Cognitive reserve. In *Neuropsychologia* 47 (10), pp. 2015–2028. DOI: 10.1016/j.neuropsychologia.2009.03.004.

Stern, Yaakov (2012): Cognitive reserve in ageing and Alzheimer's disease. In *The Lancet. Neurology* 11 (11), pp. 1006–1012. DOI: 10.1016/S1474-4422(12)70191-6.

Stolk, Ronald P.; Rosmalen, Judith G. M.; Postma, Dirkje S.; Boer, Rudolf A. de; Navis, Gerjan; Slaets, Joris P. J. et al. (2008): Universal risk factors for multifactorial diseases. In *European journal of epidemiology* 23 (1), pp. 67–74. DOI: 10.1007/s10654-007-9204-4.

Strandhagen, Elisabeth; Berg, Christina; Lissner, Lauren; Nunez, Leyla; Rosengren, Annika; Torén, Kjell; Thelle, Dag S. (2010): Selection bias in a population survey with registry linkage: potential effect on socioeconomic gradient in cardiovascular risk. In *European journal of epidemiology* 25 (3), pp. 163–172. DOI: 10.1007/s10654-010-9427-7.

Sudlow, Cathie; Gallacher, John; Allen, Naomi; Beral, Valerie; Burton, Paul; Danesh, John et al. (2015): UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. In *PLoS medicine* 12 (3), e1001779. DOI: 10.1371/journal.pmed.1001779.

Sun, Hong; Saeedi, Pouya; Karuranga, Suvi; Pinkepank, Moritz; Ogurtsova, Katherine; Duncan, Bruce B. et al. (2022): IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. In *Diabetes Research and Clinical Practice* 183, p. 109119. DOI: 10.1016/j.diabres.2021.109119.

Tang, Eugene Y. H.; Harrison, Stephanie L.; Errington, Linda; Gordon, Mark F.; Visser, Pieter Jelle; Novak, Gerald et al. (2015): Current Developments in Dementia Risk Prediction Modelling: An Updated Systematic Review. In *PloS one* 10 (9), e0136181. DOI: 10.1371/journal.pone.0136181.

Tang, Xingyao; Zhao, Wei; Lu, Ming; Zhang, Xin; Zhang, Ping; Xin, Zhong et al. (2021): Relationship between Central Obesity and the incidence of Cognitive Impairment and Dementia from Cohort Studies Involving 5,060,687 Participants. In *Neuroscience and biobehavioral reviews* 130, pp. 301–313. DOI: 10.1016/j.neubiorev.2021.08.028.

Tapiainen, V.; Hartikainen, S.; Taipale, H.; Tiihonen, J.; Tolppanen, A.-M. (2017): Hospital-treated mental and behavioral disorders and risk of Alzheimer's disease: A nationwide nested case-control study. In *Eur. psychiatr.* 43, pp. 92–98. DOI: 10.1016/j.eurpsy.2017.02.486.

Taylor, Donald H.; Østbye, Truls; Langa, Kenneth M.; Weir, David; Plassman, Brenda L. (2009): The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. In *JAD* 17 (4), pp. 807–815. DOI: 10.3233/JAD-2009-1099.

The Lancet (2022): Lecanemab for Alzheimer's disease: tempering hype and hope. In *Lancet (London, England)* 400 (10367), p. 1899. DOI: 10.1016/S0140-6736(22)02480-1.

Thesmar, David; Sraer, David; Pinheiro, Lisa; Dadson, Nick; Veliche, Razvan; Greenberg, Paul (2019): Combining the Power of Artificial Intelligence with the Richness of Healthcare Claims Data: Opportunities and Challenges. In *PharmacoEconomics* 37 (6), pp. 745–752. DOI: 10.1007/s40273-019-00777-6.

Tingley, Dustin; Yamamoto, Teppei; Hirose, Kentaro; Keele, Luke; Imai, Kosuke (2014): mediation : R Package for Causal Mediation Analysis. In *J. Stat. Soft.* 59 (5). DOI: 10.18637/jss.v059.i05.

Tripepi, Giovanni; Jager, Kitty J.; Dekker, Friedo W.; Zoccali, Carmine (2010): Selection bias and information bias in clinical research. In *Nephron. Clinical practice* 115 (2), c94-9. DOI: 10.1159/000312871.

Tutz, Gerhard; Schmid, Matthias (2016): Modeling Discrete Time-To-Event Data. Cham: Springer International Publishing AG (Springer Series in Statistics Ser). Available online at https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=4557236.

Uusitupa, Matti; Khan, Tauseef A.; Viguiliouk, Effie; Kahleova, Hana; Rivellese, Angela A.; Hermansen, Kjeld et al. (2019): Prevention of Type 2 Diabetes by Lifestyle Changes: A Systematic Review and Meta-Analysis. In *Nutrients* 11 (11), p. 2611. DOI: 10.3390/nu11112611.

van Calster, Ben; McLernon, David J.; van Smeden, Maarten; Wynants, Laure; Steyerberg, Ewout W. (2019): Calibration: the Achilles heel of predictive analytics. In *BMC medicine* 17 (1), p. 230. DOI: 10.1186/s12916-019-1466-7.

van der Lee, Sven J.; Wolters, Frank J.; Ikram, M. Kamran; Hofman, Albert; Ikram, M. Arfan; Amin, Najaf; van Duijn, Cornelia M. (2018): The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. In *The Lancet. Neurology* 17 (5), pp. 434–444. DOI: 10.1016/S1474-4422(18)30053-X.

van Dyck, Christopher H.; Swanson, Chad J.; Aisen, Paul; Bateman, Randall J.; Chen, Christopher; Gee, Michelle et al. (2023): Lecanemab in Early Alzheimer's Disease. In *The New England journal of medicine* 388 (1), pp. 9–21. DOI: 10.1056/NEJMoa2212948.

Veronese, Nicola; Facchini, Silvia; Stubbs, Brendon; Luchini, Claudio; Solmi, Marco; Manzato, Enzo et al. (2017): Weight loss is associated with improvements in cognitive function among overweight and

obese people: A systematic review and meta-analysis. In *Neuroscience and biobehavioral reviews* 72, pp. 87–94. DOI: 10.1016/j.neubiorev.2016.11.017.

Wagner, Christoph J.; Dintsios, Charalabos Markos; Metzger, Florian G.; L'Hoest, Helmut; Marschall, Ursula; Stollenwerk, Bjoern; Stock, Stephanie (2018): Longterm persistence and nonrecurrence of depression treatment in Germany: a four-year retrospective follow-up using linked claims data. In *International Journal of Methods in Psychiatric Research* 27 (2), e1607. DOI: 10.1002/mpr.1607.

Wagner, Cornelia; Carmeli, Cristian; Jackisch, Josephine; Kivimäki, Mika; van der Linden, Bernadette W. A.; Cullati, Stéphane; Chiolero, Arnaud (2024): Life course epidemiology and public health. In *The Lancet. Public health* 9 (4), e261-e269. DOI: 10.1016/S2468-2667(24)00018-5.

Walker, K. Z.; O'Dea, K.; Gomez, M.; Girgis, S.; Colagiuri, R. (2010): Diet and exercise in the prevention of diabetes. In *Journal of human nutrition and dietetics : the official journal of the British Dietetic Association* 23 (4), pp. 344–352. DOI: 10.1111/j.1365-277X.2010.01061.x.

Wee, Jason; Sukudom, Sara; Bhat, Saiuj; Marklund, Matti; Peiris, Niridu Jude; Hoyos, Camilla M. et al. (2023): The relationship between midlife dyslipidemia and lifetime incidence of dementia: A systematic review and meta-analysis of cohort studies. In *Alzheimer's & dementia (Amsterdam, Netherlands)* 15 (1), e12395. DOI: 10.1002/dad2.12395.

Wei, Jingkai; Hu, Yirui; Zhang, Li; Hao, Qiang; Yang, Ruowei; Lu, Haidong et al. (2017): Hearing Impairment, Mild Cognitive Impairment, and Dementia: A Meta-Analysis of Cohort Studies. In *Dementia and geriatric cognitive disorders extra* 7 (3), pp. 440–452. DOI: 10.1159/000485178.

Weißbach, Rafael; Dörre, Achim; Wied, Dominik; Doblhammer, Gabriele; Fink, Anne (2024): Left-truncated health insurance claims data: theoretical review and empirical application. In *AStA Adv Stat Anal* 108 (1), pp. 31–68. DOI: 10.1007/s10182-023-00471-1.

Weißbach, Rafael; Kim, Yongdai; Dörre, Achim; Fink, Anne; Doblhammer, Gabriele (2021): Left-censored dementia incidences in estimating cohort effects. In *Lifetime data analysis* 27 (1), pp. 38–63. DOI: 10.1007/s10985-020-09505-1.

Wicke, Felix Sebastian; Glushan, Anastasiya; Schubert, Ingrid; Köster, Ingrid; Lübeck, Robert; Hammer, Marc et al. (2019): Performance of the adapted Diabetes Complications Severity Index translated to ICD-10. In *The American journal of managed care* 25 (2), e45-e49.

Wilkinson, Tim; Schnier, Christian; Bush, Kathryn; Rannikmäe, Kristiina; Henshall, David E.; Lerpiniere, Chris et al. (2019): Identifying dementia outcomes in UK Biobank: a validation study of primary care, hospital admissions and mortality data. In *European journal of epidemiology* 34 (6), pp. 557–565. DOI: 10.1007/s10654-019-00499-1.

Wimo, Anders; Seeher, Katrin; Cataldi, Rodrigo; Cyhlarova, Eva; Dielemann, Joseph L.; Frisell, Oskar et al. (2023): The worldwide costs of dementia in 2019. In *Alzheimer's & dementia : the journal of the Alzheimer's Association* 19 (7), pp. 2865–2873. DOI: 10.1002/alz.12901.

Wolters, Frank J.; Chibnik, Lori B.; Waziry, Reem; Anderson, Roy; Berr, Claudine; Beiser, Alexa et al. (2020): Twenty-seven-year time trends in dementia incidence in Europe and the United States: The Alzheimer Cohorts Consortium. In *Neurology* 95 (5), e519-e531. DOI: 10.1212/WNL.0000000000010022.

World Health Organization (2004): International statistical classification of diseases and related health problems(ICD-10). Available online at https://www.who.int/standards/classifications/classification-of-diseases.

World Health Organization (2021): Global Status Report on the Public Health Response to Dementia. 1st ed. Geneva: World Health Organization. Available online at https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=30478223.

World Health Organization (2022): International Classification of Diseases, Eleventh Revision (ICD-11). Available online at https://icd.who.int/browse11/lm/en.

World Health Organization (2023): Dementia. fact sheets. Available online at https://www.who.int/news-room/fact-sheets/detail/dementia, updated on 3/15/2023, checked on 9/9/2024.

World Health Organization (2024): The top 10 causes of death. Available online at https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death, updated on 8/7/2024, checked on 9/30/2024.

Wu, Jia-Jia; Wang, Hui-Xin; Yao, Wu; Yan, Zhen; Pei, Jin-Jing (2020): Late-life depression and the risk of dementia in 14 countries: a 10-year follow-up study from the Survey of Health, Ageing and Retirement in Europe. In *Journal of Affective Disorders* 274, pp. 671–677. DOI: 10.1016/j.jad.2020.05.059.

Wu, Qiong; Tchetgen Tchetgen, Eric J.; Osypuk, Theresa; Weuve, Jennifer; White, Kellee; Mujahid, Mahasin; Glymour, M. Maria (2015): Estimating the cognitive effects of prevalent diabetes, recent onset diabetes, and the duration of diabetes among older adults. In *Dementia and geriatric cognitive disorders* 39 (3-4), pp. 239–249. DOI: 10.1159/000368654.

Wyss-Coray, Tony (2016): Ageing, neurodegeneration and brain rejuvenation. In *Nature* 539 (7628), pp. 180–186. DOI: 10.1038/nature20411.

Xu, Weili; Qiu, Chengxuan; Gatz, Margaret; Pedersen, Nancy L.; Johansson, Boo; Fratiglioni, Laura (2009): Mid- and late-life diabetes in relation to the risk of dementia: a population-based twin study. In *Diabetes* 58 (1), pp. 71–77. DOI: 10.2337/db08-0586.

Xue, Mei; Xu, Wei; Ou, Ya-Nan; Cao, Xi-Peng; Tan, Meng-Shan; Tan, Lan; Yu, Jin-Tai (2019): Diabetes mellitus and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 144 prospective studies. In *Ageing research reviews* 55, p. 100944. DOI: 10.1016/j.arr.2019.100944.

Yaffe, Kristine; Falvey, Cherie; Hamilton, Nathan; Schwartz, Ann V.; Simonsick, Eleanor M.; Satterfield, Suzanne et al. (2012): Diabetes, glucose control, and 9-year cognitive decline among older adults without dementia. In *Archives of neurology* 69 (9), pp. 1170–1175. DOI: 10.1001/archneurol.2012.1117.

Yasuno, Fumihiko; Minami, Hiroyuki; Hattori, Hideyuki (2020): Interaction effect of Alzheimer's disease pathology and education, occupation, and socioeconomic status as a proxy for cognitive reserve on cognitive performance: in vivo positron emission tomography study. In *Psychogeriatrics : the official journal of the Japanese Psychogeriatric Society* 20 (5), pp. 585–593. DOI: 10.1111/psyg.12552.

Young, Bessie Ann; Lin, Elizabeth; Korff, Michael von; Simon, Greg; Ciechanowski, Paul; Ludman, Evette J. et al. (2008): Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. In *The American journal of managed care* 14 (1), pp. 15–23.

Yu, Ruan-Ching; Proctor, Danielle; Soni, Janvi; Pikett, Liam; Livingston, Gill; Lewis, Glyn et al. (2024): Adult-onset hearing loss and incident cognitive impairment and dementia - A systematic review and meta-analysis of cohort studies. In *Ageing research reviews* 98, p. 102346. DOI: 10.1016/j.arr.2024.102346.

Zhang, Tianqi; Shaw, Marnie; Cherbuin, Nicolas (2022): Association between Type 2 Diabetes Mellitus and Brain Atrophy: A Meta-Analysis. In *Diabetes & metabolism journal* 46 (5), pp. 781–802. DOI: 10.4093/dmj.2021.0189.

Zhang, Tong; Yu, Bin (2005): Boosting with early stopping: Convergence and consistency. In *Ann. Statist.* 33 (4). DOI: 10.1214/009053605000000255.

Zhang, Xiaoyu; Tian, Qiuyue; Di Liu; Geng, Tao; Xu, Xizhu; Ge, Siqi et al. (2020): Causal association of circulating cholesterol levels with dementia: a mendelian randomization meta-analysis. In *Translational psychiatry* 10 (1), p. 145. DOI: 10.1038/s41398-020-0822-x.

Zheng, Bang; Su, Bowen; Price, Geraint; Tzoulaki, Ioanna; Ahmadi-Abhari, Sara; Middleton, Lefkos (2021): Glycemic Control, Diabetic Complications, and Risk of Dementia in Patients With Diabetes: Results From a Large U.K. Cohort Study. In *Diabetes care* 44 (7), pp. 1556–1563. DOI: 10.2337/dc20-2850.

Zoungas, Sophia; Woodward, Mark; Li, Qiang; Cooper, Mark E.; Hamet, Pavel; Harrap, Stephen et al. (2014): Impact of age, age at diagnosis and duration of diabetes on the risk of macrovascular and microvascular complications and death in type 2 diabetes. In *Diabetologia* 57 (12), pp. 2465–2474. DOI: 10.1007/s00125-014-3369-7.

9. ORIGINAL PUBLICATIONS

Study 1

Reinke, C., Doblhammer, G., Schmid, M., Welchowski, T. (2022). Dementia risk predictions from German claims data using methods of machine learning. *Alzheimer's & Dementia*. 19 (2), 477-486.

https://doi.org/10.1002/alz.12663

Study 2

Reinke, C. (2024). The effect of diabetes in the multifaceted relationship between education and cognitive function. *BMC Public Health.* 24, 2584.

https://doi.org/10.1186/s12889-024-20156-x

Study 3

Reinke, C., Buchmann, N., Fink, A., Tegeler, C., Demuth, I., Doblhammer, G. (2022). Diabetes duration and the risk of dementia: a cohort study based on German health claims data. *Age and Ageing*, 51 (1).

https://doi.org/10.1093/ageing/afab231

Study 4

Heser, K., Fink, A., **Reinke, C.**, Wagner, M., Doblhammer, G. (2020) The temporal association between incident late-life depression and incident dementia. *Acta psychiatrica Scandinavica*. 142 (5), 402-412.

https://doi.org/10.1111/acps.13220

9.1. Study 1

Received: 12 November 2021 | Revised: 21 February 2022 | Accepted: 25 February 2022

DOI: 10.1002/alz.12663

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

FEATURED ARTICLE

Dementia risk predictions from German claims data using methods of machine learning

Constantin Reinke[1] | **Gabriele Doblhammer**[1,2] | **Matthias Schmid**[2,3] | **Thomas Welchowski**[3]

[1] Institute for Sociology and Demography, University of Rostock, Rostock, Germany

[2] German Center for Neurodegenerative Diseases, Bonn, Germany

[3] Institute of Medical Biometry, Informatics and Epidemiology (IMBIE), Medical Faculty, University of Bonn, Bonn, Germany

Correspondence
Constantin Reinke, Institute for Sociology and Demography, University of Rostock, Ulmenstr. 69, 18057 Rostock, Germany.
Email: constantin.reinke@uni-rostock.de

Abstract

Introduction: We examined whether German claims data are suitable for dementia risk prediction, how machine learning (ML) compares to classical regression, and what the important predictors for dementia risk are.

Methods: We analyzed data from the largest German health insurance company, including 117,895 dementia-free people age 65+. Follow-up was 10 years. Predictors were: 23 age-related diseases, 212 medical prescriptions, 87 surgery codes, as well as age and sex. Statistical methods included logistic regression (LR), gradient boosting (GBM), and random forests (RFs).

Results: Discriminatory power was moderate for LR (C-statistic = 0.714; 95% confidence interval [CI] = 0.708–0.720) and GBM (C-statistic = 0.707; 95% CI = 0.700–0.713) and lower for RF (C-statistic = 0.636; 95% CI = 0.628–0.643). GBM had the best model calibration. We identified antipsychotic medications and cerebrovascular disease but also a less-established specific antibacterial medical prescription as important predictors.

Discussion: Our models from German claims data have acceptable accuracy and may provide cost-effective decision support for early dementia screening.

KEYWORDS
calibration, dementia, discrimination, Germany, health claims data, machine learning, risk factors

1 | BACKGROUND

There are currently ≈47 million people worldwide living with dementia. This number is expected to increase to 78 million by 2030 and to 132 million by 2050.[1] In Germany, 1.6 million people are presently living with dementia, with an expected increase to 2.7 million by 2050.[2] Dementia creates high costs for society and the health care system, which increase significantly as the disease progresses.[3]

Because dementia is still incurable, prevention is the best strategy to delay its onset and to slow progression, with the goal of reducing the burden of dementia on those affected and on the health care system. For effective prevention, it is crucial to identify modifiable risk factors and to detect cognitive decline at an early stage prior to manifestation, even better, before the onset.

To achieve this, numerous dementia risk-prediction models have been developed with different target populations and outcomes. Most models predict the risk of late-life dementia for non-demented people, but there are also midlife risk models and models for the conversion from mild cognitive impairment (MCI) to dementia. Outcomes include Alzheimer's disease (AD), other dementia subtypes, combinations of subtypes, and all-cause dementia. However, reviews stress the need for further models in different populations[4,5] and point out that more recent and larger data sets are needed to overcome the lack of diversity in previous studies.[6] To address this limitation, routinely collected

health care data has come more into focus in recent years.[7–11] These data are typically cost-effective because they do not need to be collected separately. Typical routinely collected data are electronic health records (EHRs) as well as administrative health claims data. Although the introduction of EHR in Germany has already been legally decreed, the implementation, collection, and availability of the data have not yet been established,[12] making them hardly accessible for research due to data protection regulations. In the German health care system, one of the largest health care systems worldwide, it is mandatory to participate in a health insurance fund, and nearly 90% of the population is covered by public health insurance. Although administrative health claims data are used primarily for the purpose of billing for health care services, these claims contain a large amount of clinically relevant information. They also include sociodemographic data, as well as all diagnoses made by physicians, operations performed, and medications (prescriptions filled). Therefore, claims data are used increasingly in public health research.

An illustrative example of the use of routinely collected data for disease prediction is the QRISK, a cardiovascular disease prediction risk score that complements the prevailing Framingham Risk Score.[13]

Early detection of dementia would require a life-course approach, with major vascular risk factors already becoming prevalent during midlife[14] and the diagnosis relying strongly on the medical history of the patient.[15] Health claims data have the potential to provide such long-term information.

Previous studies have used a number of statistical methods to predict dementia risk. In recent years the most common approaches have been logistic and Cox regression as well as an increasing number of machine learning (ML) techniques.[6] ML algorithms are well suited for the analysis of data sets with a large amount of information, as they usually contain automatic variable selection mechanisms and can include non-linear associations as well as complex interactions between variables.[16] Several studies have used ML algorithms for risk prediction with administrative claims data. For example, a recent study compared different ML methods and traditional models to predict heart failure outcomes, achieving the best performance with gradient-boosting models and logistic regression.[17] Another study found an improved accuracy of cardiovascular risk prediction using ML and electronic medical records.[18] Nori and colleagues identified incident dementia by applying ML algorithms to an administrative claims data set of privately insured individuals in the United States.[9] Moreover, logistic regression has been used to predict dementia diagnosis from administrative claims.[10] These studies identified neurological and psychological disorders and psychoactive medications as predictors with the greatest impact on dementia risk. Discriminatory power, as measured by the area under the curve and the concordance index, ranged between 0.63 and 0.76.

In this context we examined the following research questions: First, are German claims data a suitable data basis for individual dementia-risk prediction? Second, how do ML methods compare to classical regression methods in terms of predicting dementia risk? Third, which features are important predictors of dementia risk, and can new features be identified in addition to established risk factors?

RESEARCH IN CONTEXT

1. **Systematic review:** An increasing number of dementia risk- prediction models have been developed and the most common methodologies are machine learning (ML) and traditional regression methods. Despite the increasing availability of routinely collected health data in Germany, dementia risk predictions using these data are still rare.
2. **Interpretation:** German claims data are suitable for dementia risk prediction. We found moderate prediction accuracy for logistic regression and gradient boosting. In addition to some well-known dementia-related features, we identified the pharmacological subgroup of macrolides, lincosamides, and streptogramins (ATC-code: J01F) as an important predictor for dementia.
3. **Future directions:** Dementia risk-prediction models from German claims data may be useful in implementing cost-effective decision-support tools for early dementia screening. Data-driven approaches with claims data have the potential to identify new features or pathways affecting the risk of dementia.

HIGHLIGHTS

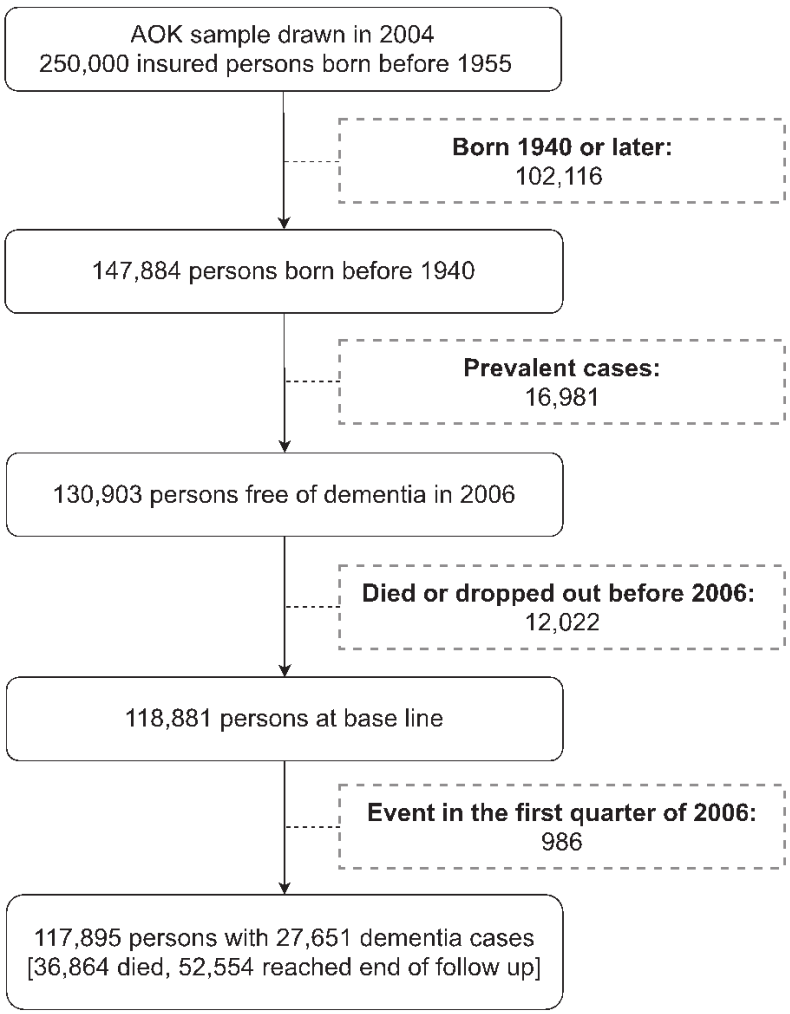
- Gradient boosting machine (GBM) showed the best model calibration.
- The prediction accuracy of GBM was comparable to that of classical logistic regression.
- Model accuracies were comparable and partly better than other studies using claims.
- Selected antibacterial medical prescriptions were an important predictor for dementia.

2 | MATERIAL AND METHODS

2.1 | Data

We used an age-stratified random sample of 250,000 persons insured in the largest German health insurance company "Allgemeine Ortskrankenkasse" (AOK). The sample was drawn in 2004 and included people born before 1955 with a follow-up to 2015. The AOK covers almost 30% of the German population and is representative of the German population aged 65+ in terms of mortality (Figure S1). The data contained the following information from the inpatient and outpatient sector: Diagnoses based on the International Classification of Diseases, Tenth Revision (ICD-10), all medical prescriptions according to the Anatomical Therapeutic Chemical (ATC) Classification

FIGURE 1 Selection of study cohort



System, surgeries based on the German procedure classification (OPS, www.dimdi.de/dynamic/en/classifications/ops/, accessed: November 1, 2021) an adaptation of the International Classification of Procedures in Medicine (ICPM), as well as sex, age, and time of death. All medical information was recorded on a quarterly basis. The data were anonymized claims data and did not require ethical review or patient consent.

2.2 | Study design

To predict incident dementia, we selected people born before 1940 because dementia before age 65 is extremely rare. Because we were interested in newly diagnosed dementia cases only, we excluded all people with prevalent dementia in 2004 or 2005. In the quarter of a dementia diagnosis, we assigned all predictors the values of the previous quarter because the exact timing of diagnoses, prescriptions, and surgeries or the chronological order of occurrence was unknown. Hence, we excluded all persons with an incident dementia diagnosis in the first quarter of 2006 (Figure 1), and the study period began with the first quarter of 2006 and ended with the last quarter of 2015. We assembled a training population by drawing a 60% random sample from our study cohort, stratified by dementia status (dementia diagnosis vs no dementia diagnosis). The remaining individuals were split into validation and test groups of equal size.

2.3 | Outcome

Our outcome was a binary variable that indicated a validated incident dementia diagnosis. Dementia was defined by ICD-10 codes (Table S1). To address the problem of false-positive diagnosis, we applied a two-stage validation strategy[19] (Description S1).

2.4 | Predictors

We included 23 major age-related diseases[20] and risk factors for dementia according to the factors reported by the Lancet Commission[21] (see Table S1 for the respective ICD codes). In addition, we included all medical prescriptions coded by the German version of

the ATC classification system on the 3rd level (pharmacological subgroup), and all surgeries based on three-digit classes from Chapter 5 of the OPS. Further predictors were sex and age at baseline.

All predictors except age and sex were included as time-dependent binary "ever" variables, with the value 1 from the first occurrence of a particular code onward and zero otherwise.

2.5 | Statistical analysis

The data were structured quarterly, with one observation per person and quarter. Therefore, we defined a discrete time process starting in the first quarter of 2006, with time intervals given by quarters, resulting in a data structure with binary outcome.[22] Accordingly, time measurements ranged between 1 and 40, referring to the number of quarters observed for each individual. To predict dementia risk, we built prediction models using logistic regression, gradient boosting machines (GBMs), and random forests (RFs). The input data for these models consisted of one row per quarter, thereby allowing for time-dependent predictors and accounting for right-censoring ("discrete hazard models," cf.[22]) We excluded predictors with near-zero-variance using the nearZeroVar function from,[23] and predictors with fewer than five observations per cell in a cross-tabulation with dementia. Surgery codes 502 and 583 were excluded due to collinearity. In total, we included 324 features: 212 ATC codes, 87 OPS codes, 23 diagnoses, and age and sex.

As a benchmark model, we considered a logistic regression model (including all predictors) that was fitted to the combined training and validation data, and the process time was included as a categorical predictor.

We used the R package xgboost version 1.1.1.1[24] to train a GBM with a learning rate of 0.01 and a maximum of 10,000 iterations on the training data. The algorithm was stopped if there was no improvement in log-loss in the last 100 iterations evaluated on the validation data.[24] We used a grid search on the validation data to find optimal values for the parameters *max depth* and *min child weight*. Finally, we fitted a GBM with the optimal parameter values to the combined training and validation data.

We used the R package ranger version 0.12.1[25] to train RF with ntree = 1000 and ntree = 10,000 trees on the training data. All other parameters were set to default. Because logistic regression and GBM clearly outperformed RF, we did not perform any further parameter tuning here.

Because a detailed description of the GBM and RF methodology is beyond the scope of this applied study, we included our source code in the supplements (Supplementary_source_code). This code provides detailed information about our analysis and preprocessing methods.

2.6 | Model evaluation

We evaluated the performance of our prediction models in terms of accuracy, discriminatory power, and calibration.

The overall accuracy was evaluated by the integrated prediction error, an adaptation of the Brier score, which is based on weighted quadratic differences between predicted and observed survival functions.[22] We calculated the integrated prediction error using the R package discSurv[26]; 95% confidence intervals (95% CIs) were calculated using 5000 bootstrap replications from the test data.

To evaluate the discriminatory power, we calculated a time-independent version of the concordance index (C statistic).[27] As with the prediction error, we used the discSurv package and calculated 95% CIs using bootstrapping.

Because a strong discriminatory power is not sufficient to assess a model for clinical usability,[28] we additionally examined calibration plots (to graphically assess model calibration) and calculated an intercept and slope to test whether the predicted risks were systematically overestimated or underestimated.[29]

To extract the explainable information from the GBM model, we identified the 20 most influential features for prediction, using a permutation approach to calculate a relative importance score for each predictor.[30] To present explainable information for logistic regression, we report the corresponding odds ratios, ranked by absolute z values. Because GBM clearly outperformed RF, we did not calculated variable importance for RF.

3 | RESULTS

3.1 | Study cohort

Our study cohort consisted of 117,895 individuals, and we observed 27,651 incident dementia cases. During the study period, 63,864 individuals died (details in Figure S2) and 52,554 reached the end of follow-up (Figure 1). The training data consisted of 70,737 people, and the validation data included 23,579 people, as did the test data. At baseline, the mean age was 74.8 years (SD = 6.6), and the mean age at dementia diagnosis was 82.3 years (SD = 6.3); 38% of the individuals were men (See Table 1). The mean follow-up time was 17.8 quarters (SD = 11.3).

3.2 | Model evaluation

Logistic regression indicated the strongest discriminatory power on the test data, with a C statistic of 0.714 (95% CI = 0.708–0.720), closely followed by GBM with a C statistic of 0.707 (95% CI = 0.700–0.713). Although the CIs overlapped (Summary in Table S2), pairwise differences in C values were different from zero (mean = 0.007, 95% CI = 0.005–0.009). The discriminatory power of RF (ntree = 1000) was considerably lower, with a C statistic of 0.636 (95% CI = 0.628–0.643). The same model-ranking sequence appeared for the integrated prediction error, where the lowest error was found for the logistic regression model (0.044, 95% CI = 0.044–0.045), followed by GBM (0.046, 95% CI = 0.046–0.047) and RF (0.105, 95% CI = 0.104–0.107). Looking at the calibration plots (Figure 2), the GBM appeared to be a well-calibrated model with an intercept near zero (0.097) and a slope close

TABLE 1 Cohort characteristics at baseline

Training	Number	%	2.5% Quantile	97.5% Quantile
N	70,737			
Age at baseline, mean (SD)	72.8 (6.6)		64	89
Men	26,997	38.2		
Dementia cases (not at baseline)	16,522	23.4		
Antipsychotics (ATC: N05A)	3,528	5.0		
Cerebrovascular disease	14,528	20.5		
Anti-dementia drugs (N06D)	966	1.4		
Depression	13,547	19.2		
Parkinson disease	1,489	2.1		
Injuries to the head	3,734	5.3		
Antidepressants (N06A)	8,729	12.3		
Cardiomyopathy and heart failure	17,075	24.1		
Test				
N	23,579			
Age at baseline, mean (SD)	72.9 (6.6)		64	88
Men	9,063	38.4		
Dementia cases (not at baseline)	5,506	23.4		
Antipsychotics (N05A)	1,090	4.6		
Cerebrovascular disease	4,900	20.8		
Anti-dementia drugs (N06D)	330	1.4		
Depression	4,458	18.9		
Parkinson disease	479	2.0		
Injuries to the head	1,200	5.1		
Antidepressants (N06A)	2,848	12.1		
Cardiomyopathy and heart failure	5,794	24.6		
Total				
N	117,895			
Age at baseline, mean (SD)	72.8 (6.6)		64	89
Men	45,038	38.2		
Dementia cases (not at baseline)	27,651	23.5		
Antipsychotics (N05A)	5,827	4.9		
Cerebrovascular disease	24,242	20.6		
Anti-dementia drugs (N06D)	1,639	1.4		
Depression	22,548	19.1		
Parkinson disease	2,491	2.1		
Injuries to the head	6,176	5.2		
Antidepressants (N06A)	14,601	12.4		
Cardiomyopathy and heart failure	28,575	24.2		

Table 1 shows the baseline cohort characteristics for the training, test, and full data sets. In addition to age and sex, we report only the 10 most influential predictors in terms of variable importance (Figure 2). In total, we included 212 ATC codes, 87 OPS codes, 23 diagnoses, and age and sex. Source: AOK data 2004-2015, own calculations.

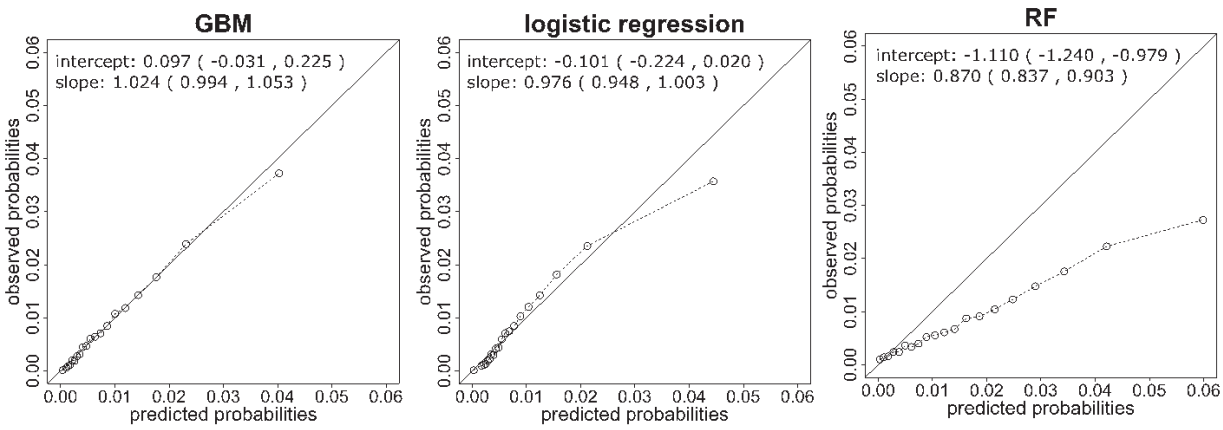


FIGURE 2 Model calibration on test data for GBM (gradient boosting machine), logistic regression, and RF (random forest). Intercepts and slopes were calculated by logistic calibration (95% confidence intervals in parentheses). The models showed weaker performance in the higher risk segments compared to the medium and lower segments

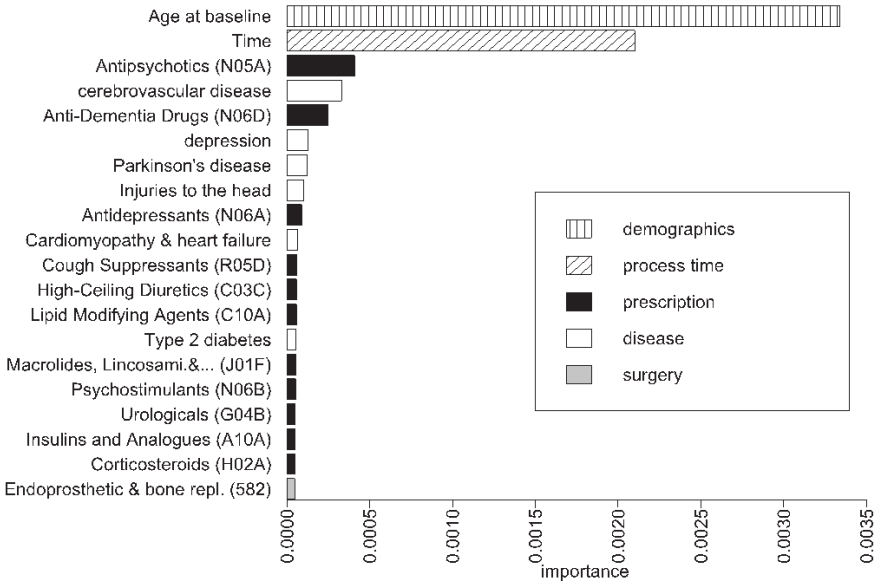


FIGURE 3 Variable importance of 20 of the most influential features from the GBM (gradient boosting machine) model. *Source*: AOK data 2004-2015, own calculations

to 1 (1.024). The logistic regression model was calibrated slightly worse (intercept: −0.101; slope: 0.976) and the RF showed substantial issues with calibration (intercept: −1.110; slope: 0.870). Predictive performance of the RF with ntree = 10,000 was very similar to the performance of RF with ntree = 1000 (Table S2). The performance of RF may be improved further by hyper-parameter tuning.

3.3 | Most important predictors

By far the most important predictor in the GBM model was age at baseline (Figure 3). The most important medical prescriptions were antipsychotics (N05A), anti-dementia drugs (N06D), and antidepressants (N06A), all of which were among the top 10 features. Most interestingly, 11 of the top 20 features were medical prescriptions; among these some medications associated with diseases linked to a high risk of dementia (insulin for diabetes mellitus, diuretics for high blood pressure). However, we also found medication that has not been described in the context of dementia prediction, such as macrolides, lincosamides and streptogramins (J01F), and medication that has been described as being protective in previous studies, for example, corticosteroids (H02A).

Among the diseases, cerebrovascular diseases were most important, followed by depression, Parkinson disease, and injuries of the

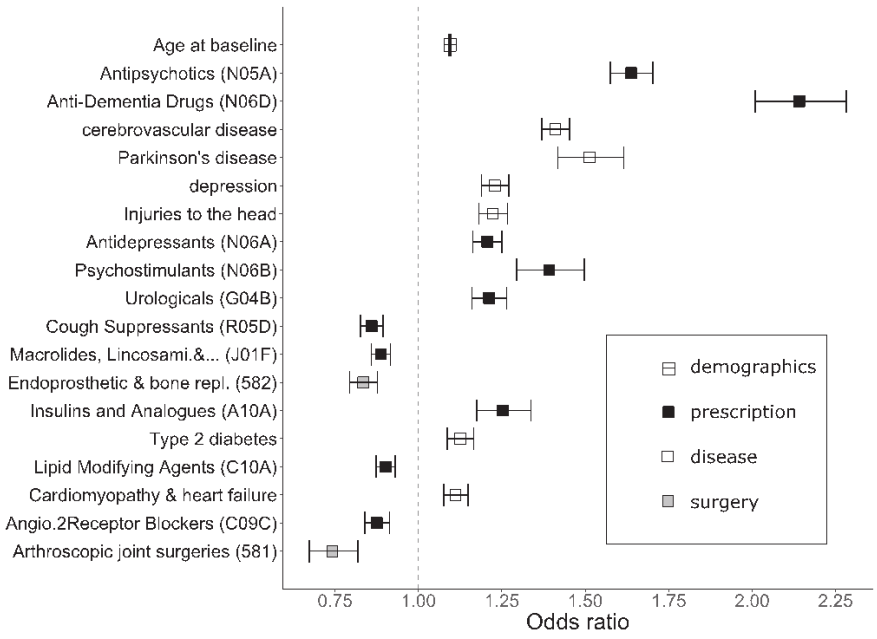


FIGURE 4 Odds ratios with 95% confidence intervals from logistic regression according to the highest absolute z values (except time). *Source*: AOK data 2004-2015, own calculations

head. In addition, cardiomyopathy and heart failure and diabetes mellitus were among the top 20. For surgeries, endoprosthetic and bone replacement was an important feature.

The odds ratios (Figure 4) show a slightly different ranking than the variable importance, however, most variables are identical. Angiotensin-2 receptor blockers and arthroscopic joint surgeries occurred here instead of high-ceiling diuretics and corticosteroids.

4 | DISCUSSION

We applied and compared classical and ML methods to develop dementia risk-prediction models using German health claims data. To our knowledge, this is the first study using German claims data for dementia-risk prediction.

With C statistics higher than 0.70, both logistic regression and GBM indicated acceptable discriminatory power. When compared to clinical studies, which used information on biomarkers, cognitive test scores, or laboratory results, this discriminatory power is somewhat lower.[4] This may be because much of this information is more closely related to diagnostic approaches for dementia than individual medical history. For example, cognitive tests are usually part of the diagnostic process. Nevertheless, our results are comparable to or partly better than the results of other studies, which used claims data only.[9,10,31]

The ML methods did not outperform classical logistic regression in terms of discriminatory power, as indicated by our test on pairwise differences. However, GBM showed a better calibration. Regarding interpretability, odds ratios obtained from logistic regression are usually easier to interpret than variable importance measures obtained from ML methods. Specifically, odds ratios come along with a sign (indicating risks vs protective effects) and are well suited for explanatory purposes (provided that the model has been specified correctly). On the other hand, variable importance measure provides more insight in the *predictive* ability of variables, incorporating effects due to non-linearity and variable selection.

Most of the features identified among the top 20 are consistent with previous results about important risk and predictive factors for dementia. By far the most important feature for predicting dementia in general and in our study is age. Also outstanding are antipsychotics (N05A), which are used frequently to treat a range of psychiatric symptoms[32] as well as for the pre-dementia stage of MCI.[33,34] Although the evidence for the benefit of antidementia drugs (N06D) in the pre-dementia stage is inconclusive,[35] they are widely used in clinical practice.[36] Note that the high importance of anti-dementia drugs may also indicate a high proportion of undiagnosed dementia cases, especially for individuals in the phase of conversion from MCI to dementia, which is a phase of high uncertainty.[37]

The high importance of cerebrovascular disease for incident dementia (feature ranks fourth after age and antipsychotics) may indicate a high proportion of vascular dementia. A further important cardiovascular-related feature is cardiomyopathy and heart failure, which underscores the importance of the association of cardiovascular diseases and dementia. Type 2 diabetes and insulin (A10A) can also be included in this category. Considering that drug utilization can be an indicator of diseases,[38] high-ceiling diuretics (C03C) may indicate hypertension. Neurodegenerative risk factors such as head injury, which is associated with an increased risk of dementia[39] and Parkinson disease,[40] are ranked seventh and eighth. Late-life depression is a

known risk factor for dementia,[41] but may also be a prodromal symptom of dementia.[42] There is also evidence for an association between antidepressants and an increased risk of dementia,[43] but the prescriptions of antidepressants could also indicate undiagnosed depression. The only surgery code in the top 20 was endoprosthetic and bone replacement (OPS:582). There is evidence that although the risk of dementia was increased in the quarter of endoprosthetic and bone replacement surgery, it was lower in the postoperative period than for those without surgery.[44] Psychostimulants contain medication for the treatment of attention deficit/hyperactivity disorder (ADHD), which is associated with an increased risk of dementia.[45] Urologicals are partly associated with cognitive impairment[46] and are prescribed commonly for urinary incontinence, which has been identified as a predictor of dementia in previous studies.[9,10]

Features with evidence of a protective association with dementia risk were also included in the top 20 most important features, for example, cough suppressants,[47] lipid-modifying agents,[48] and corticosteroids.[49] However, we also found some less-established dementia-related features, such as medical prescriptions of the macrolides, lincosamides, and streptogramins (ATC:J01F). The association of inflammation and bacterial infections with cognitive decline and dementia are well established,[50] whereas the role of antibacterial medications and the risk of dementia is largely unclear.[51] Antibiotics, for one thing may damage the microbiome in the gut leading to a higher dementia risk.[52] Then again they may reduce inflammation leading to a lower risk.[53]

Although the discriminatory power of our approach suggests a rather moderate relevance for direct identification of dementia cases in clinical practice, it highlights the potential of health claims data as supporting tools for prediction of dementia risk. Dementia-risk predictions from claims data can be used as a cost-effective indication for the need for additional dementia screening. The combination of claims data with further information may improve the discriminatory power[17] and increase relevance in clinical practice.[54] One reason that ML did not perform better than logistic regression in our study could be the large number of 0-1 coded features (only age was continuous). By including additional information, such as laboratory values or test parameters, ML may perform better than logistic regression, as associations with dementia risk are not necessarily linear.

The use of data-driven ML methods in routinely collected data can be an important contribution to identifying or better understanding known pathways and new risk and preventive factors for dementia (such as the feature macrolides, lincosamides and streptogramins (J01F), which we found). The ability of ML methods to include complex interactions between several features may be key to the study of multimorbidity and effects of polypharmacy.[55] These strengths of ML methods are especially important against the background of increasing availability of high-dimensional data in health care.

4.1 | Strengths and limitations

The large longitudinal and population-based data containing information from the inpatient and outpatient sector as well as those living in nursing homes is representative of the older German population, which adds to the strength of the study conducted. Because the data are collected routinely in a standardized fashion, problems such as sample selection bias, attrition, and recall bias are less relevant than in other data sources. We applied an established strategy to validate dementia diagnosis. Three different statistical methods were used, which included more than 300 characteristics as time-dependent predictors. In addition to discriminatory power, we also explored the calibration of the models using a large internal test set.

We acknowledge some limitations in this study. Although claims data offer some advantages, it is important to note that this type of data is used primarily for billing purposes and the information available was generated for health care utilization only. For example, diagnoses can be identified only for people who went to a doctor, which limits the generalizability of our results. In addition, dementia may be underreported, especially in the early pre-clinical stages. Furthermore, a large proportion of dementia diagnoses are unspecific, which does not allow for an accurate distinction between dementia subtypes. The information about medical prescriptions is limited to the collection of the drug (redemption of the prescription), without any information about the actual intake. Moreover, the proportion of persons with low socioeconomic status is higher in the AOK than in other statutory health insurance companies and also in comparison with private health insurance companies.[56] Although these differences could influence both morbidity and the utilization of health care services, they can be explained partly by the different age structure of the AOK population, which is older than the German population. At the same age, the difference in the social structure of the AOK population is larger in younger age groups than in older.[56] Age-specific mortality[57] and age-specific prevalence and incidence of dementia in AOK data are similar to those shown for the total German population.[19] The diagnoses included here are limited to a manual selection of 23 age-related conditions, which seems counterintuitive in the context of a data-driven approach. However, consideration of all ICD-10 diagnosis codes at the disease group level (three digits) might not be accurate in medical terms, and inclusion of all diagnosis codes (more than 13,000) would exceed our computational resources and would lead to very sparse data. Because our models are limited to Germany and were validated with data from the same source, the generalizability of our results needs further investigation. Future studies should also investigate the effects of competing events (eg, death before dementia diagnosis).

5 | CONCLUSION

Our results from routinely collected claims data are not suitable for making diagnoses or replacing established tests in clinical practice, but they may be useful as an additional measure for risk detection. Specifically, they may be useful in implementing decision support for early dementia screening in a cost-effective manner if health care providers could continuously update physicians on current risk predictions. Ideally, the cornerstones of dementia diagnosis such as clinical assessment, laboratory testing, and imaging[15] should be combined into one data repository. Undoubtedly, more research on different types of

health data is needed before any real benefit of such an approach at the public health level can be determined. Our results may also be relevant to dementia prevention research in order to identify new features or pathways that influence dementia risk. The combination of claims data with data-driven approaches may serve as a starting point for further research into the largely unknown association between dementia and characteristics such as certain types of antibacterial medical prescriptions identified in this study.

ACKNOWLEDGEMENT

The authors are grateful to the Scientific Research Institute of the Allgemeine Ortskrankenkassen (AOK), (German: Wissenschaftliches Institut der AOK, WIdO), for providing the data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Constantin Reinke https://orcid.org/0000-0003-3228-1794

REFERENCES

1. Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M. World Alzheimer Report 2015 - The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International, 2015.
2. Alzheimer Europe. Dementia in EuropeYearbook 2019: Estimating the Prevalence of Dementia in Europe. Luxembourg: Alzheimer Europe; 2019.
3. El-Hayek YH, Wiley RE, Khoury CP, et al. Tip of the Iceberg: assessing the Global Socioeconomic Costs of Alzheimer's Disease and related dementias and strategic implications for stakeholders. *J Alzheimers Dis*. 2019;70(2):323–341.
4. Hou X-H, Feng L, Zhang C, Cao X-P, Tan L, Yu J-T. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry*. 2019;90(4):373–379.
5. Tang EYH, Harrison SL, Errington L, et al. Current developments in dementia risk prediction modelling: an updated systematic review. *PLoS One*. 2015;10(9):e0136181.
6. Goerdten J, Čukić I, Danso SO, Carrière I, Muniz-Terrera G. Statistical methods for dementia risk prediction and recommendations for future work: a systematic review. *Alzheimers Dementia (New York, N Y)*. 2019;5:563–569.
7. Ben Miled Z, Haas K, Black CM, et al. Predicting dementia with routine care EMR data. *Artif Intell Med*. 2020;102:101771.
8. Ford E, Greenslade N, Paudyal P, et al. Predicting dementia from primary care records: a systematic review and meta-analysis. *PLoS One*. 2018;13(3):e0194735.
9. Nori VS, Hane CA, Martin DC, Kravetz AD, Sanghavi DM. Identifying incident dementia by applying machine learning to a very large administrative claims dataset. *PLoS One*. 2019;14(7):e0203246.
10. Albrecht JS, Hanna M, Kim D, Perfetto EM. Predicting diagnosis of Alzheimer's disease and related dementias using administrative claims. *J Manag Care Spec*. 2018;24(11):1138–1145.
11. Jain S, Rosenbaum PR, Reiter JG, et al. Using Medicare claims in identifying Alzheimer's disease and related dementias. *Alzheimers Dementia*. 2020;17(3):515–524.
12. Pohlmann S, Kunz A, Ose D, et al. Digitalizing health services by implementing a personal electronic health record in Germany: qualitative analysis of fundamental prerequisites from the perspective of selected experts. *J Med Internet Res*. 2020;22(1):e15102.
13. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007;335(7611):136.
14. Reitz C, Brayne C, Mayeux R. Epidemiology of Alzheimer disease. *Nat rev Neurol*. 2011;7(3):137–152.
15. Gauthier S, Rosa-Neto P, Morais JA, Webster C. World Alzheimer Report 2021 - Journey Through the Diagnosis of Dementia. London: Alzheimer's Disease International, 2021.
16. Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Statist Sci*. 2001;16(3):199–231.
17. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA network open*. 2020;3(1):e1918962.
18. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;12(4):e0174944.
19. Doblhammer G, Fink A, Fritze T. Short-term trends in dementia prevalence in Germany between the years 2007 and 2009. *Alzheimers Dementia*. 2015;11(3):291–299.
20. Doblhammer G, Barth A. Prevalence of morbidity at extreme old age in germany: an observational study using health claims data. *J Am Geriatr Soc*. 2018;66(7):1262–1268.
21. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet North Am Ed*. 2020;396(10248):413–446.
22. Tutz G, Schmid M. *Modeling Discrete Time-to-Event Data*, Springer Ser. Statistics. 1, Cham: Springer International Publishing; 2016:1–247.
23. Kuhn M. Building Predictive Models in R Using the caret Package. *J Stat Soft*. 2008;28(5):1–26.
24. Zhang T, Yu B. Boosting with early stopping: convergence and consistency. *Ann Statist*. 2005;33(4):1538–1579.
25. Wright MN, Ziegler A. Ranger A fast implementation of random forests for high dimensional data in C++ and R. *J Stat Soft*. 2017;77(1):1–17.
26. Welchowski T, Schmid M. discSurv: Discrete Time Survival Analysis. R package version 1.1.4; 2019. https://cran.r-project.org/web/packages/discSurv/discSurv.pdf
27. Schmid M, Tutz G, Welchowski T. Discrimination measures for discrete time-to-event predictions. *Econom Stat*. 2018;7:153–164.
28. van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC med*. 2019;17(1):230.
29. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. Medical decision making an international journal of the Society for. *Med Decis Making*. 1993;13(1):49–58.
30. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5–32.
31. Fukunishi H, Nishiyama M, Luo Y, Kubo M, Kobayashi Y. Alzheimer-type dementia prediction by sparse logistic regression using claim data. *Comput Methods Programs Biomed*. 2020;196:105582.
32. Gareri P, Segura-García C, Manfredi VGL, et al. Use of atypical antipsychotics in the elderly: a clinical review. *Clin Interv Aging*. 2014;9:1363–1373.
33. Söderlund ME, Guajardo ME, Cavagna M, et al. Prediction of antipsychotics use in patients with mild cognitive impairment and dementia. *Alzheimers Dementia*. 2020;16(S8). https://doi.org/10.1002/alz.046645
34. Dhikav V. Predictors of antipsychotic usage among patients with dementias and mild cognitive impairment. *Alzheimers Dementia*. 2020;16(S7). https://doi.org/10.1002/alz.043398

35. Fink HA, Jutkowitz E, McCarten JR, et al. Pharmacologic interventions to prevent cognitive decline, mild cognitive impairment, and clinical Alzheimer-type dementia: a systematic review. *Ann Intern Med*. 2018;168(1):39–51.
36. Bertens D, Vos S, Kehoe P, Wolf H, et al. Use of mild cognitive impairment and prodromal AD/MCI due to AD in clinical care: a European survey. *Alzheimer's Res Ther*. 2019;11(1):74.
37. Bruscoli M, Lovestone S. Is MCI really just early dementia? A systematic review of conversion studies. *Int Psychogeriatr*. 2004;16(2):129–140.
38. Chini F, Pezzotti P, Orzella L, Borgia P, Guasticchi G. Can we use the pharmacy data to estimate the prevalence of chronic conditions? A comparison of multiple data sources. *BMC public health*. 2011;11:688.
39. Li Y, Li Y, Li X, Zhang S, et al. Head injury as a risk factor for dementia and Alzheimer's disease: a systematic review and meta-analysis of 32 observational studies. *PLoS One*. 2017;12(1):e0169650.
40. Aarsland D, Kurz MW. The epidemiology of dementia associated with Parkinson's disease. *Brain Pathol*. 2010;20(3):633–639.
41. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. *Br J Psychiatry*. 2013;202(5):329–335.
42. Heser K, Fink A, Reinke C, Wagner M, Doblhammer G. The temporal association between incident late-life depression and incident dementia. *Acta Psychiatr Scand*. 2020;142(5):402–412.
43. Moraros J, Nwankwo C, Patten SB, Mousseau DD. The association of antidepressant drug usage with cognitive impairment or dementia, including Alzheimer disease: a systematic review and meta-analysis. *Depress Anxiety*. 2017;34(3):217–226.
44. Teipel SJ, Fritze T, Ellenrieder M, Haenisch B, Mittelmeier W, Doblhammer G. Association of joint replacement surgery with incident dementia diagnosis in German claims data. *Int Psychogeriatr*. 2018;30(9):1375–1383.
45. Tzeng N-S, Chung C-H, Lin F-H, et al. Risk of dementia in adults with ADHD: a nationwide, population-based cohort study in Taiwan. *J Atten Disord*. 2019;23(9):995–1006.
46. Kim YJ, Tae BS, Bae JH. Cognitive function and urologic medications for lower urinary tract symptoms. *Int Neurourol J*. 2020;24(3):231–240.
47. Hwang T-J, Chen J-J, Lin Y-T, Chan H-Y. Dextromethorphan for the treatment of agitation in dementia: a pilot study. *Alzheimers Dementia*. 2020;16(S9). https://doi.org/10.1002/alz.045791
48. Ancelin M-L, Carrière I, Barberger-Gateau P, et al. Lipid lowering agents, cognitive decline, and dementia: the three-city study. *J Alzheimers Dis*. 2012;30(3):629–637.
49. Nerius M, Haenisch B, Gomm W, Doblhammer G, Schneider A. Glucocorticoid therapy is associated with a lower risk of dementia. *J Alzheimers Dis*. 2020;73(1):175–183.
50. Muzambi R, Bhaskaran K, Brayne C, Davidson JA, Smeeth L, Warren-Gash C. Common bacterial infections and risk of dementia or cognitive decline: a systematic review. *J Alzheimers Dis*. 2020;76(4):1609–1626.
51. Kern DM, Cepeda MS, Lovestone S, Seabrook GR. Aiding the discovery of new treatments for dementia by uncovering unknown benefits of existing medications. *Alzheimers Dementia (New York, N Y)*. 2019;5:862–870.
52. Angelucci F, Cechova K, Amlerova J, Hort J. Antibiotics, gut microbiota, and Alzheimer's disease. *J Neuroinflammation*. 2019;16(1):108.
53. Balducci C, Forloni G. Doxycycline for Alzheimer's disease: fighting β-amyloid oligomers and neuroinflammation. *Front pharmacol*. 2019;10:738.
54. Rao A, Miller B, Kulman T, Pinho P, Aggarwal NT. Novel application of digital dementia phenotyping and risk classification for insurance and longevity risk modeling. *Alzheimers Dementia*. 2020;16(S10). https://doi.org/10.1002/alz.044372
55. Hassaine A, Canoy D, Solares JRA, et al. Learning multimorbidity patterns from electronic health records using non-negative matrix factorisation. *J Biomed Inform*. 2020;112:103606.
56. Epping J, Geyer S, Eberhard S, Tetzlaff J. Völlig unterschiedlich oder doch recht ähnlich? Die soziodemografische Struktur der AOK Niedersachsen im Vergleich zur niedersächsischen und bundesweiten Allgemein- und Erwerbsbevölkerung. *Gesundheitswesen*. 2021;83(S 02):77–86.
57. Nerius M, Fink A, Doblhammer G. Parkinson's disease in Germany: prevalence and incidence based on health claims data. *Acta Neurol Scand*. 2017;136(5):386–392.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Reinke C, Doblhammer G, Schmid M, Welchowski T. Dementia risk predictions from German claims data using methods of machine learning. *Alzheimer's Dement*. 2022;1–10. https://doi.org/10.1002/alz.12663

9.2. Study 2

Reinke *BMC Public Health* (2024) 24:2584
https://doi.org/10.1186/s12889-024-20156-x

BMC Public Health

RESEARCH Open Access

The effect of diabetes in the multifaceted relationship between education and cognitive function

Constantin Reinke[1*]

Abstract

Background Education has been shown to be positively associated with cognitive performance. However, the pathways via lifestyle-related disease through which education is related to cognitive performance have not been sufficiently explored. Diabetes is an important lifestyle-related disease with increasing prevalence worldwide. Low education is associated with an increased risk of developing diabetes, while diabetes may also lead to a deterioration in cognitive performance. This study aims to explore if the associations between education and cognitive function is mediated by the diabetes status among older adults.

Methods The data utilized in this study were derived from the first two waves of the Dutch Lifelines Cohort Study (2006–2015). The analyzed sample included 26,131 individuals aged 50 years or above at baseline. The baseline assessment included measurements of educational attainment (exposure) and the potential mediator diabetes. The outcome of cognitive function was assessed using age-standardized reaction times from the psychomotor function and attention tasks, as measured by the Cogstate Brief Battery. The Cogstate Brief Battery was only conducted at the follow-up assessment, not at the baseline assessment. Faster reaction times correspond to higher cognitive performance. The study employed linear and logistic regression models, in addition to a causal mediation approach which estimated the average causal mediation effect (ACME).

Results Higher education was associated with a lower risk of diabetes (b= -0.1976, 95%CI= -0.3354; -0.0597) compared to low or middle education as well as with faster reaction times (b= -0.2023, 95%CI= -0.2246; -0.1798), implying better cognitive function. Diabetes was associated with slower reaction times (b = 0.0617, 95%CI = 0.0162; 0.1072). Most importantly, the mediation approach identified a significant indirect effect of education on cognitive function via the diabetes status (ACME= -0.00061, 95%CI= -0.00142; -0.00011).

Discussion The findings emphasize the potentially importance of diabetes in explaining the role of education in promoting healthy cognitive function and mitigating the risk of cognitive decline. Early detection and treatment of diabetes may be particularly beneficial for individuals with low or middle levels of education in order to maintain good levels of cognitive function.

Keywords Cognitive function, Education, Diabetes, Mediation, Population study

*Correspondence:
Constantin Reinke
constantin.reinke@uni-rostock.de
[1]Institute for Sociology and Demography, University of Rostock, Ulmenstr. 69, 18057 Rostock, Germany

Introduction

The doubling of the population aged 60 years and older by 2050 [1], in conjunction with the continued aging of the global population, will result in a significant increase in the global challenge of cognitive decline. The decline in cognitive function affects the individual's daily activities, resulting in a diminished quality of life and loss of independence with a high burden on caregivers and health care systems. While brain changes associated with cognitive decline are part of normal brain aging [2], certain diseases such as diabetes can accelerate neurodegeneration and also be a driver of cognitive decline [3]. In 2021, it is estimated that there are 537 million people with diabetes, with a predicted increase to 783 million by 2045 [4]. Conversely, factors such as education can compensate for or delay cognitive decline [5, 6]. In many parts of the world, an impressive expansion of education has taken place during the last decades [7], which, however, leaves those with less education at a particular risk of disease and poor cognition [6, 8].

As cognitive aging and brain-altering processes are irreversible, the best strategy for reducing the risk of cognitive decline or delaying the onset or progression to clinical manifestations such as dementia is to identify and address those risk factors that are amenable to modification. The Lancet Commission identified less education as well as diabetes as important modifiable risk factors for dementia [9]. Thus, a better understanding of the pathways and possible links of these two factors with cognitive function may contribute to potential strategies for preventing or rather delay cognitive decline and dementia.

Diabetes is associated with deficits in cognitive function [10, 11] and a higher risk of cognitive impairment [12] as well as dementia [13]. Furthermore, there is evidence for a link between diabetes and brain atrophy which leads to deficits in cognitive function [14]. High blood glucose levels and hyperglycemic events affect the brain by cerebral microvascular dysfunctions [15] and can lead to brain atrophy [16]. On the other hand, diabetes is associated with higher risks of a series of cardiovascular diseases [17], which are known to be the main drivers of cognitive impairment and vascular dementia [18]. However, cognitive function can also be affected by high blood glucose levels and hyperglycemic events due to more complex pathways including oxidative stress and neuroinflammation [19]. Prior meta-analyses have indicated that diabetes exerts disparate effects on various domains of cognitive function, notably affecting psychomotoric function and attention [10, 11].

The association between education and cognitive function is well researched [20] and established by the concept of Cognitive Reserve [21]. The accumulation of education and experience of occupational complexity over a lifetime strengthens resilience to the pathology of cognitive decline due to age-related brain changes and delays the symptoms of cognitive decline or the clinical manifestation of dementia. In addition, there is evidence that education may also interact with the level of tau protein accumulation in the brain and its role in cognitive function [22]. Education has been identified as the most important proxy measure for cognitive reserve [23] and this is true for both individuals with and without diabetes as shown by the similar relationship between cognitive reserve level and a variety of executive cognitive function scores independent from the diabetes status [23].

The risk of diabetes is lower in individuals with higher levels of education [24, 25]. Although other determinants of socioeconomic status have been studied and associated with the diabetes risk, education has been the most frequently and consistently associated indicator [26]. The underlying mechanism for the relationship between education and diabetes is not fully understood, but factors related to lifestyle and healthy behaviors appear to play a crucial role. In particular, BMI has been identified as an important factor [27, 28].

It is reasonable to conclude that the most influential factor in the association of education and cognition is the direct link, as evidenced by the aforementioned connections of intellectual stimulation, such as the theory of cognitive reserve [20, 21]. Nevertheless, the pathways connecting education and cognitive functions are multifaceted and remain incompletely understood. Lower education is linked to a higher risk of lifestyle-related diseases [29], which are also associated with cognitive decline. This suggests that individuals with lower education levels may be more vulnerable to cognitive decline due to these conditions. Beside diabetes, also other lifestyle related diseases like cardiovascular diseases [30], vascular diseases [31], the number of chronic diseases [32] or obesity [33] are associated with worse cognitive function and represent possible determinates to play a role in the link between education and cognitive function. Among these diseases, diabetes is of particular interest because it is a modifiable risk factor for cognition in multiple ways. Prevention of diabetes and good glucose management in people with diabetes are both important to reduce the risk of cognitive decline [34, 35].

However, the interplay of education and diabetes and the consequence for cognitive function is still less researched and remain largely unknown.

Kowall & Rathmann examined the combined effects of education and diabetes on cognitive performance using longitudinal data from more than 27 countries from the SHARE project [36]. The authors found that people with diabetes had worse cognitive performance than people without diabetes, and that people with diabetes had even worse cognitive performance if they had lower levels

of education. However, because the authors found no interaction effect between education and diabetes, they concluded that the effects were additive. A small retrospective case-control study including 1537 individuals from Japan examined the pathway between socioeconomic status and dementia by evaluating lifestyle-related disease as potential mediators [37]. However, the authors did not find a significant association between educational attainment and the risk of diabetes in their data, so a conclusion about the role of diabetes as a potential mediator is limited.

Because the pathways through which education is related to cognitive performance have not been well studied, this study addressed the question of whether some of the association between education and cognitive function may operate through the diabetes status. The hypothesis is that the diabetes status partly mediates the association between education and cognitive function.

Materials and methods

Data

Analyses were conducted using data from the Dutch Lifelines Cohort Study. *Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 167,729 persons living in the North of the Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics* [38]. The large dataset includes information on physical examinations, biological samples, cognitive tests and a comprehensive questionnaire. Data collection was conducted between 2006 and 2013 for the baseline assessments and between 2014 and 2015 for the second assessment. Lifelines was conducted in accordance with the guidelines of the Declaration of Helsinki and has been approved by Medical ethical committee of the University Medical Center Groningen (The Netherlands) under number 2007/152. All participants signed an informed consent form.

Study design & sample

The Lifelines cohort includes 152,860 individuals aged 18 years or older at baseline. Of these, 111,959 participated in the second assessment. All individuals younger than 50 years at baseline were excluded, as were individuals with missing data on outcome (no Cogstate examination), exposure (education), mediator (diabetes), or confounders. See Fig. 2. The final sample consisted of 26,131 individuals. To set up a study design with a causal time order, the variables for the exposure, the mediator and confounders were built by information from the baseline assessment. As the analysis only included individuals aged 50 or older at baseline, it can be assumed that their highest educational attainment was achieved well before the mediators and confounders were measured. Since the Cogstate examination was not conducted at baseline, the outcome measure was taken from the second assessment only. As a result, it was not possible to investigate the change in cognitive function over time or to adjust for baseline cognitive function.

Data availability

Data may be obtained from a third party and is not publicly available. Researchers may apply to use the Lifelines data used in this study. For information on how to request Lifelines data and terms of use are available on their website at (https://www.lifelines.nl/researcher/how-to-apply).

Cognitive function measure (outcome)

Individuals cognitive function was measured by tasks from the Cogstate Brief Battery at the second assessment only. The Cogstate Brief Battery is a validated computer

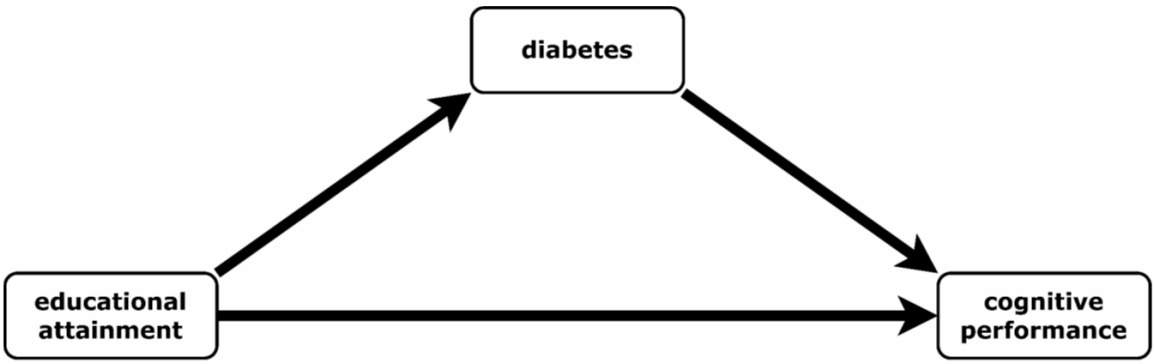


Fig. 1 Hypothesized relationship between outcome (cognitive performance), mediator (diabetes) and exposure (educational attainment), source: Own illustration

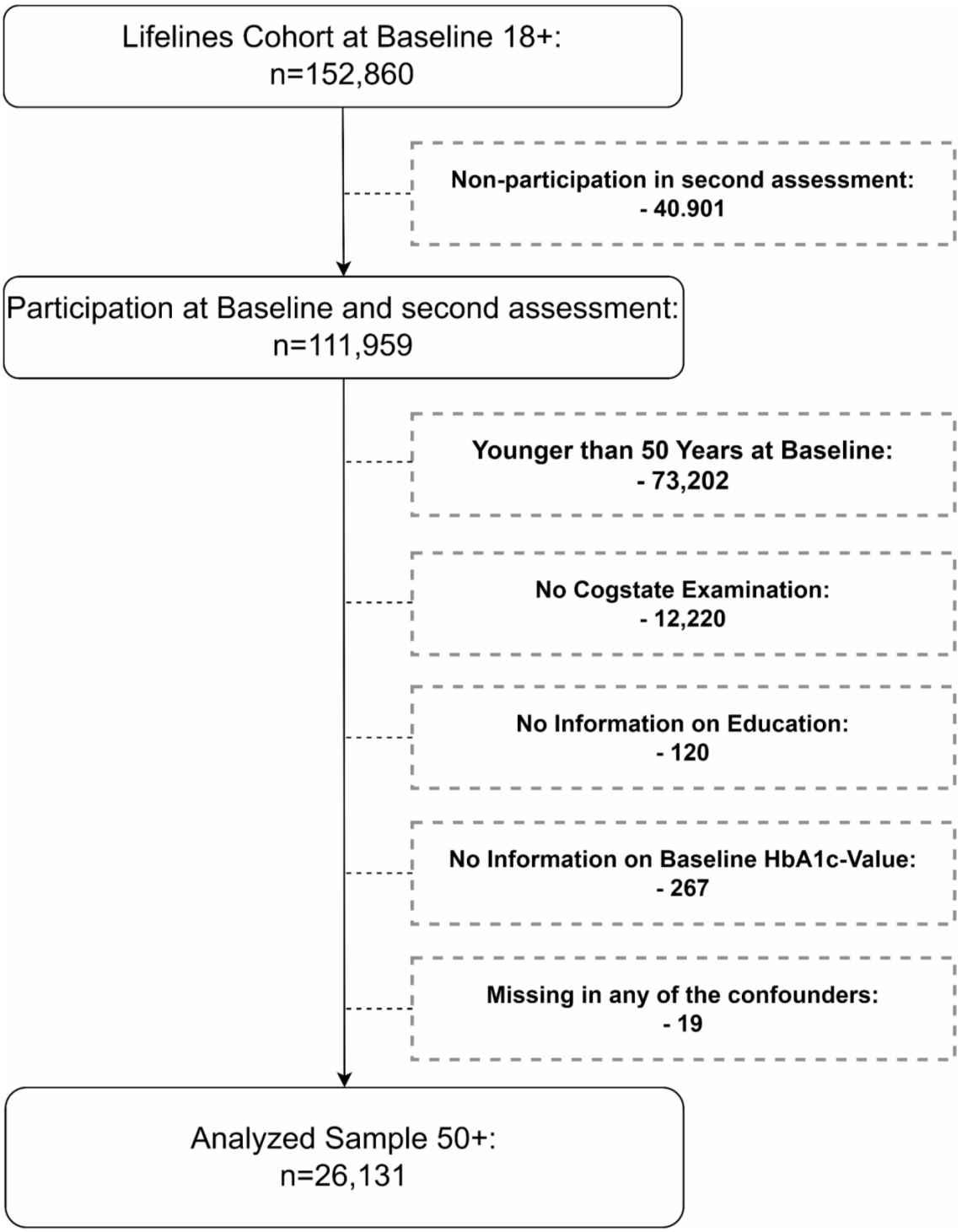


Fig. 2 Selection of the study cohort, source: lifelines data 2006–2015, own calculation

based cognitive assessment [39]. It has been used in previous studies to detect mild cognitive impairment and cognitive impairment in Alzheimer's disease [40] and also in the context of the Lifelines cohort [41]. The battery includes four different tasks to measure the cognitive domains of psychomotoric function (detection test), attention (identification test), visual learning (one card learning test), and working memory (one back test). Outcomes of each task were reaction time and accuracy.

Previous studies point out, that the associations of diabetes and cognitive function varied between the domains of cognitive function [10, 11]. There is evidence that primarily the domains of psychomotoric function and attention were affected by diabetes. Therefore, a composite score from these two domains was built by the detection and identification tasks of the Cogstate Brief Battery, similar to Maruff and colleagues [40]. In order to achieve this, the log-10-transformed (closer to a normal distribution) reaction time (milliseconds) of both tests were z-standardized in 5-year age groups (from age group 50 upwards) and then summed up (Distribution: Supplementary Figure S1). Accordingly, a positive value indicates a higher (slower) reaction time compared to the respective age group, representing poorer cognitive function. Conversely, a negative value displays a lower (faster) reaction time, implying better cognitive function.

Educational attainment (exposure)

Individuals educational attainment was defined by the highest obtained degree at baseline. The information was self-reported through a questionnaire with the following possible responses: no education (1); primary education (2); lower or preparatory secondary vocational education (3); junior general secondary education (4); secondary vocational education or work-based learning pathway (5); senior general secondary education, pre-university secondary education (6); higher vocational education (7); university education (8); other (9). We categorized the education into two categories low-middle (1–6) and high education (7–8). Individuals who reported "other" were assigned to one of the two categories in a further step (see: https://wiki.lifelines.nl/doku.php?id=educational_attainment).

Diabetes status (mediator)

Diabetes was defined as the presence of at least one of the following conditions at baseline: Self-reported diagnosis of diabetes, baseline HbA1c ≥ 6.5%, fasting plasma glucose ≥ 7 mmol/L, random plasma glucose ≥ 11.1 mmol/L or use of any medical diabetes treatment.

Control variables

The statistical models were adjusted for age, sex, physical activity, obesity, smoking history, income and a set of comorbidities: depression, hypertension, stroke, heart failure and high cholesterol. The models on cognitive function were also controlled for the accuracy (number correct responses/number total response) of the tasks from the Cogstate Brief Battery. All confounders were measured at the baseline assessment.

Physical activity was defined as any amount of vigorous physical activity per week. So, the variable is classified into physical active vs. not physical active persons.

Obesity was defined by a body mass index of 30 or higher.

The smoking history was classified by using the cumulative risk measure of packyears, where one pack year implies smoking 20 cigarettes (or an equivalent number of other smoking derivatives) per day for one year. The variable was categorized into: Never smokers, persons with equal or less pack years than the median of the ever smokers, persons with more pack years than the median of the ever smokers, no answer/pack years not calculable.

Income was measured by the following question: "what is your net income per month? (if you share a household, include the net income of your partner(s))". The variable is classified into the following categories: Lower than 1500€, 1500€ □2500€, Over 2500€, Don't know/no answer.

All comorbidities were measured by self-report (had the person ever had the condition).

Statistical analysis

To explore the associations between education, diabetes status and cognitive function linear and logistic regression models were used. A causal mediation approach [42] was used to test whether there was an indirect effect of education (**E**xposure) on cognitive function (**O**utcome) through diabetes status (**M**ediator) controlled for the **c**onfounders. This approach which based on the idea or structural equation modeling estimates the average causal mediation effect (ACME) or indirect effect. The approach included several steps: First, two statistical models were fitted separately, one to model the mediator (1) and one to model the outcome (2). In a second step, model parameters for outcome and mediator were simulated from their sampling distributions. Third, potential values for the mediator were simulated prior to the potential outcomes, given these simulated mediator values, in order to then calculate the causal mediation effects from the simulated values. Finally, the point estimation for the ACME (and direct effect) as well as the confidence interval was calculated from the simulated distribution. The approach is described in detail by Imai et al. [42]

$$M = \beta_0 + \beta_1 E + \sum_{i=1}^{n} \beta_{1+i} c_i + \varepsilon_M \quad (1)$$

$$O = \alpha_0 + \alpha_1 E + \alpha_2 M + \sum_{i=1}^{n} \alpha_{1+i} c_i + \varepsilon_o \quad (2)$$

In contrast to the classical framework of structural equation modeling, here the two models do not have to be linear regression models. Since the mediator (diabetes status) is a binary variable, it was estimated using logistic regression. While the outcome is a continuous variable, it was estimated by linear regression. Both models were estimated with robust standard errors. The causal mediation approach was performed using the mediation package in R [43].

To check the robustness of the results, a number of sensitivity analyses were applied. First, to test the sequential ignorability assumption that there are no unmeasured confounders of the mediator-outcome relation [44], the "medsens" function from the mediation package in R [43] was used. The function estimates a parameter which indicates the correlation of ε_M and ε_o at which the ACME would be zero. The function required a probit regression model instead of a logistic one.

As a further sensitivity analysis, a classical structural equation modeling approach with robust standard errors was used. However, in this classical framework both models have to be linear regression models, the mediator was the (continuous) HbA1c-level instead of the diabetes status otherwise the equations were identical to (1) and (2). After the simultaneously estimation of the outcome and the mediator model, the effect of education was decomposed in a direct and indirect effect and the Sobel test was used to check significance [45].

To test the robustness of the study design, the mediation analysis described above was also tested using indicators other than diabetes as potential mediators. In the same way as for diabetes, the indirect effect of education was tested via high blood pressure, high cholesterol and obesity.

In a further sensitivity analysis, the diabetes status was defined as the presence of at least one of the conditions mentioned in Sect. 2.5 at baseline or follow-up.

All analyses were performed using R 4.2.2 and Stata 17.

Results

Study cohort

Our analyzed sample included 26,131 individuals of which 18,486 had a low or middle educational attainment and 7645 had a high educational attainment, and there were 1449 diabetes cases. The mean age-standardized reaction time of the outcome measure was □0.0003 (SD = 0.8796). The mean age at baseline was 58.1 (SD = 6.6) years, with a range of 50 to 88 years. The majority of participants were between 50 and 60 years of age. The mean age was 57.9 (SD = 6.5) without and 61.1 (SD = 6.9) for persons with diabetes. The individuals are divided into 14,503 women and 11,628 men. Table 1 shows characteristics of the study cohort. When looking at the outcome, it can be noticed that the mean age-standardized reaction time for high educated people (-0.2175) is clearly lower (faster/better) than for people with low or middle education (0.0895). So, in the cognition tests individuals with high education performed faster/better compared to their age-group while low-middle educated individuals performed slower/worse. A reverse pattern was observed in diabetes. Non-diabetic individuals demonstrated faster/better cognitive performance relative to their age group than diabetic individuals.

Mediation analysis

Table 2 shows the estimated regression coeficients for the mediator model 1 (logistic) and the outcome model 2 (Ordinary Least Squares) with the corresponding confidence intervals (CI) as well as the direct and the indirect effect (ACME) of education on cognitive function. From model 1 it can be derived that higher education is associated with a lower risk of diabetes compared to low-middle education, indicated by the negative regression coeficient (-0.1976, $p = 0.005$). The regression coefficient (-0.2023, $p < 0.001$) in model 2 shows a significant association between education and cognitive function, implying that those with higher education have a lower reaction time to the outcome measure, and therefore better cognitive function than those with low or middle education. The regression coeficient for diabetes in model 2 is 0.0617 ($p = 0.008$). This demonstrates that individuals with diabetes have a significantly higher reaction time to the outcome measure and therefore worse cognitive function than individuals without diabetes. The average mediation effect of education trough diabetes was □0.00061 and the direct effect of education was □0.20247. This results in a total effect of -0.20307. In particular, the coeficient indicates that individuals with higher levels of education, in comparison to those with low-to-middle levels of education, completed the examined tasks, on average, 0.20307 standard deviation faster (better cognitive function) than the average in their respective age group. All these effects were significantly different from zero. However, the ratio of the indirect effect to the total effect implies that the indirect effect contributes less than 1% to the total effect. The average mediation effect (-0.00061) represents the difference in the effect of education on cognitive function through the mediator. In other words, it is the total effect minus the direct effect.

Table 1 Study cohort characteristics

Variable	Persons at baseline	Persons with diabetes at baseline (%)	psychomo-tor function and attention tasks* (SD)
Age group			
50–54	9569	285 (2.98)	0.0097 (0.8923)
55–59	6259	312 (4.98)	-0.0061 (0.8812)
60–64	5627	384 (6.82)	-0.0056 (0.8800)
65–69	3201	304 (9.50)	-0.0047 (0.8512)
70–74	1123	120 (10.69)	-0.0130 (0.8525)
75–79	287	34 (11.85)	-0.0034 (0.8384)
80+	65	9 (13.85)	-0.0064 (0.8466)
Educational attainment			
low-middle	18,486	1130 (6.12)	0.0895 (0.9109)
High	7645	318 (4.16)	-0.2175 (0.7559)
Diabetes			
No	24,682		-0.0070 (0.8769)
yes	1449		0.1137 (0.8715)
Sex			
Female	14,503	693 (4.78)	0.0325 (0.8790)
male	11,628	756 (6.50)	-0.0410 (0.8787)
Physical activity			
not active	17,930	1160 (6.47)	0.0445 (0.8907)
Active	8201	289 (3.52)	-0.0983 (0.8468)
Obesity			
No	21,920	839 (3.83)	-0.0098 (0.8775)
yes	4211	610 (14.49)	0.0495 (0.8890)
Smoking history			
never smoker	8842	415 (4.69)	0.0196 (0.8906)
equal or less pack years than the median of ever smokers	8020	337 (4.20)	-0.0118 (0.8781)
more pack years than the median of ever smokers	7929	587 (7.40)	-0.0343 (0.8595)
no answer/pack years not calculable	1340	110 (8.21)	0.1377 (0.9170)
Household income per month			
lower than 1500€ per month	2589	185 (7.15)	0.1221 (0.9149)
1500€ – 2500€ per month	7279	463 (6.36)	0.0552 (0.8811)
over 2500€ per month	11,670	503 (4.31)	-0.1123 (0.8342)
don't know/no answer	4593	298 (6.49)	0.1274 (0.9331)
Depression			
No	23,618	1280 (5.42)	-0.0037 (0.8783)
yes	2513	169 (6.73)	0.0316 (0.8913)
Stroke			
No	25,836	1413 (5.47)	-0.0022 (0.8790)
yes	295	36 (12.2)	0.1621 (0.9246)
Hypertension			
No	17,641	626 (3.55)	-0.0102 (0.8791)
yes	8490	823 (9.69)	0.0203 (0.8804)
Heart failure			
No	25,816	1401 (5.43)	-0.0014 (0.8785)
yes	315	48 (15.24)	0.0888 (0.9601)
High cholesterol			
No	19,965	690 (3.46)	-0.0078 (0.8764)
yes	6,166	759 (12.31)	0.0240 (0.8896)
Total	**26,131**	**1449 (5.55)**	**-0.0003 (0.8796)**

Source: lifelines data 2006–2015, own calculation

** Mean age-standardized reaction time [log10-transformed milliseconds]*

Table 2 Results of regression models for the mediator and the outcome variable

	model 1 * (mediator)		model 2 † (outcome)
dependent variable:	diabetes status		cognitive function
model type:	logistic		OLS
	Reg. coef. (95% CI)		**Reg coef. (95% CI)**
High education (Ref.: low-middle)	-0.1976 (-0.3354; -0.0597)		-0.2023 (-0.2246; -0.1798)
Diabetes (Ref.: No-Diabetes)	-		0.0617 (0.0162; 0.1072)
(Pseudo) R-squared	0.1108		0.0847
Number of observations	26,131		26,131
ACME of education (indirect effect)		-0.00061 (-0.00142; -0.00011) 0.3% of the total effect	
direct effect of education		-0.20247 (-0.22516; -0.18052) 99,7% of the total effect	
total effect of education		-0.20307 (-0.226145; -0.18077)	

regression coefficients and 95% confidents intervals & direct, indirect and total effect of education from the causal mediation analysis, source: lifelines data 2006–2015, own calculation.

** Model controlled for: Age, sex, physical activity, obesity, smoking history, income, and hypertension.*

† Model controlled for: age, sex, physical activity, obesity, smoking history, income, comorbidities, and cognition test accuracy.

The first sensitivity analysis tested the sequential ignorability assumption of the causal mediation approach. The estimated parameter ρ that would lead to an ACME of zero was 0.036 (Supplementary Figure S2). So even a weak pre-treatment confounder could render the effect insignificant.

In a second sensitivity analysis, a classical structural equation modelling approach was applied with HbA1c level instead of diabetes status as a potential mediator. The results of the Sobel test showed that HbA1c was also a significant mediator for education (Supplementary Table S1). The proportion of the indirect effect from the total effect was also less than 1%.

A further sensitivity analysis was conducted to test the robustness of the study design by using other lifestyle-related diseases as potential mediators. The results showed that there were no indirect effects (ACME not significantly different from zero) of education on cognitive function via hypertension, high cholesterol, or obesity (Supplementary Table S2 – Table S4).

The sensitivity analysis, which defined diabetes status based on baseline and follow-up information, resulted in a larger number of cases of diabetes (2013 vs. 1449). However, the results of the mediation analysis did not differ significantly from those presented in Table 2 (see Supplementary Table S5).

Discussion

The question of this study was whether diabetes partly mediates the link between educational attainment and cognitive function in individuals aged 50 years and older, using a large data set from the Netherlands. The results revealed significant positive effects of higher education on cognitive function as well as a lower risk of diabetes for higher educated individuals. The most noteworthy finding was the identification of a significant indirect effect of education on cognitive function via diabetes, although this effect was relatively small.

Although the mediating effect of diabetes on the relationship between education and cognition has not been considerably studied, an earlier study did examine the interaction effect between education and diabetes [36]. They did not find one and therefore concluded that the effects of the two risk factors were purely additive. This conclusion does not rule out that the connection may be partly a mediated association. In addition to the increased risk of developing diabetes among individuals with lower levels of education, potential reasons for the association to cognition may lie in the disparities in health literacy and adherence to diabetes therapy between the educational groups. Supporting Kowall and Rathmann, worse glycemic control which is more prevalent in lower educated people with diabetes [46, 47] and the association of worse glycemic control with cognitive dysfunction [48] may be a mechanism here. Treatment recommendations for glycemic control are challenging and include diabetes self-management by monitoring of blood glucose, use of medication as well as physical activity and nutrition/diet [49]. This health-related behaviors are linked to education [46, 50, 51]. Furthermore, the compliance of diabetes self-management decreases over time [52] and there is evidence that individuals with lower levels of education are at an increased risk of developing diabetes complications [53] which point out the possible link to cognitive function. This may give a higher potential for reducing the burden of cognitive decline in lower educated people by avoiding diabetes as well as

diabetes complications through improving diabetes self-management and adherence.

Nakahori and colleagues [37] concluded that diabetes plays a minimal role in the link between educational attainment and dementia. This was not surprising as their study showed no significant association between education and diabetes, or between education and other dementia risk factors such as smoking. However, the authors emphasized that their findings are limited to Japan. In contrast to the findings of Nakahori et al., this study confirmed the existing evidence for the connection of education and diabetes [24, 25, 31] and argued that diabetes plays a significant role in the link between education and cognitive function. It should be noted that Nakahori et al. employed dementia as outcome, rather than cognitive function. Nevertheless, both are closely related and share a significant number of risk factors. A further reason for the disparate findings may be attributed to the differing definitions of diabetes applied. While Nakahori et al. only include diagnosed cases of diabetes, this study also incorporates individual's laboratory results, which cover undiagnosed cases of diabetes. The link between undiagnosed diabetes and cognitive function appears to be particularly strong [54].

One of the main strengths of our study is its large sample size. The lifelines cohort covers about 10% of the population of the northern Netherlands. In this study more than 26,000 people were included and analyzed using information from questionnaires, measurements, and blood sample data. Central to the validity of the study is the utilization of a validated and well-established measure for assessing cognitive function. A composite score of cognitive function was constructed using tests from the Cogstate Brief Battery. This score was derived from two domains: psychomotor function and attention. There is a body of evidence indicating that these domains are associated with diabetes [10, 11]. An additional strength of our study is the comprehensive definition of diabetes, which includes both diagnosed and undiagnosed cases by incorporating the HbA1c-level from the blood sample. Furthermore, the statistical models were adjusted for a set of life style-related confounders. The sensitivity analyses concerning a further statistical approach, as well as the robustness of the study design using other possible mediators or extended definition of diabetes represent further strengths of this study.

Despite the strengths of this study, it is important to acknowledge its limitations. The Cogstate Brief Battery was not assessed at baseline; thus, it was not possible to examine changes in cognitive function in the context of the analyses. Moreover, the cognitive function at follow-up is supposed to be affected by the cognitive function at baseline; however, it was not possible to adjust the analysis for this information. In light of the findings of this study, it can be concluded that the outcome measure of cognitive function encompasses both the baseline information and the change from baseline to follow-up. Consequently, it is not possible to interpret these elements separately. It is relevant to consider this when interpreting the results of the study. It is necessary for future studies to take this issue into account in order to strengthen the validity of these findings. Considering the sensitivity analysis and the low robustness to the assumption of pre-treatment confounding, it is necessary to reflect on the results with this in mind. In particular, genetic factors [55–57] may be important, but further unobserved factors may also be pre-treatment factors that are connected to education, diabetes, and cognitive function such as e.g. the socioeconomic background in childhood [58, 59].

The multifaceted relationship between education and cognition has already been pointed out, and several lifestyle factors that are related to education, diabetes and cognition have been controlled in the model. Nevertheless, the complex relationship with nutrition could not be modelled, even if obesity was included in the statistical model. However, we did not find a significant pathway via hypertension, high cholesterol, or obesity in our statistical model which strengthens the importance of diabetes as one of the multifaceted pathways between education and cognition.

Further issues concern the analyzed sample. Population-based health surveys are typically affected by selection or response bias, which leads to a healthier study sample than in the underlying population. This is also suspected here, with a diabetes prevalence of 5.7% for the study sample (age 50+) towards 7.5% in the Netherland population aged 20–79 [60]. It is reasonable to assume that this also applies to individuals with impaired cognitive function. This may result in an underestimation of the effects in the statistical models and with it the size/proportion of the indirect effect of education on cognitive function. Moreover, the follow-up period between the baseline and second assessment was relatively short given the slow progression of diabetes and cognitive decline.

The relationship between education and cognitive abilities is well established and again evidenced by the findings of this study. Education, particularly in older age, is not a modifiable risk factor, whereas diabetes is. Thus, in the multifaceted relationship between education and cognition, diabetes represents one promising approach to modifying or preventing the risk of cognitive decline thereby counteracting the disadvantages of less education. This holds true, even if the association between education and cognitive function may be mediated by a series of factors, among them most prominently cognitive reserve.

Further research is needed to reveal how cognitive performance changes over time in the different educational groups. Additionally, the interplay between genetic predisposition, education and cognitive function should be explored.

Conclusion

This study found that people with lower levels of education were more likely to have diabetes and that diabetes was associated with poor cognitive function. Most importantly, this study is the first to demonstrate that a part of the effect of education on cognitive function runs through diabetes. While the relationship between education and cognition is multifaceted, these findings emphasize the potentially importance of diabetes in explaining the role of education in promoting healthy cognitive function and mitigating the risk of cognitive decline. Lower and middle educated people are double disadvantaged with respect to cognitive function through a higher risk for diabetes as well as a lower cognitive reserve resulting from lower education. Early detection and treatment of diabetes may be particularly beneficial for these individuals to maintain good levels of cognitive function.

Abbreviations

ACME	Average causal mediation effect
CI	Confidence interval
HbA1c	Glycated hemoglobin (A1c)
mmol/L	Millimole per liter
SD	Standard deviation
UMCG	University Medical Center Groningen

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-024-20156-x.

Supplementary Material 1

Acknowledgements

The services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participants are gratefully acknowledged. The Lifelines initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG), Groningen University and the Provinces in the North of the Netherlands (Drenthe, Friesland, Groningen).

Author contributions

This paper was written by a single author. The author was responsible for all aspects of the research process, including the study design and conception, the analysis and interpretation of data, and preparation of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Open Access funding enabled and organized by Projekt DEAL.

Data availability

Data may be obtained from a third party and is not publicly available. Researchers may apply to use the Lifelines data used in this study. For information on how to request Lifelines data and terms of use are available on their website at (https://www.lifelines.nl/researcher/how-to-apply).

Declarations

Ethics approval and consent to participate

The Lifelines study was conducted in accordance with the guidelines of the Declaration of Helsinki and has been approved by Medical ethical committee of the University Medical Center Groningen (The Netherlands) under number 2007/152. All participants signed an informed consent form.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 24 July 2024 / Accepted: 23 September 2024
Published online: 27 September 2024

References

1. World Health Organization. Ageing and health: Key facts. 2022. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health. Accessed 14 Jun 2024.
2. Ferreira LK, Busatto GF. Resting-state functional connectivity in normal brain aging. Neurosci Biobehav Rev. 2013;37:384–400. https://doi.org/10.1016/j.neubiorev.2013.01.017.
3. Verdile G, Fuller SJ, Martins RN. The role of type 2 diabetes in neurodegeneration. Neurobiol Dis. 2015;84:22–38. https://doi.org/10.1016/j.nbd.2015.04.008.
4. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract. 2022;183:109119. https://doi.org/10.1016/j.diabres.2021.109119.
5. Lenehan ME, Summers MJ, Saunders NL, Summers JJ, Vickers JC. Relationship between education and age-related cognitive decline: a review of recent research. Psychogeriatrics. 2015;15:154–62. https://doi.org/10.1111/psyg.12083.
6. Zahodne LB, Stern Y, Manly JJ. Differing effects of education on cognitive decline in diverse elders with low versus high educational attainment. Neuropsychology. 2015;29:649–57. https://doi.org/10.1037/neu0000141.
7. Lutz KCS. The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. Glob Environ Change. 2017;42:181–92. https://doi.org/10.1016/j.gloenvcha.2014.06.004.
8. Beydoun MA, Beydoun HA, Gamaldo AA, Teel A, Zonderman AB, Wang Y. Epidemiologic studies of modifiable factors associated with cognition and dementia: systematic review and meta-analysis. BMC Public Health. 2014;14:643. https://doi.org/10.1186/1471-2458-14-643.
9. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396:413–46. https://doi.org/10.1016/S0140-6736(20)30367-6.
10. Monette MCE, Baird A, Jackson DL. A meta-analysis of cognitive functioning in nondemented adults with type 2 diabetes mellitus. Can J Diabetes. 2014;38:401–8. https://doi.org/10.1016/j.jcjd.2014.01.014.
11. Palta P, Schneider ALC, Biessels GJ, Touradji P, Hill-Briggs F. Magnitude of cognitive dysfunction in adults with type 2 diabetes: a meta-analysis of six cognitive domains and the most frequently reported neuropsychological tests within domains. J Int Neuropsychol Soc. 2014;20:278–91. https://doi.org/10.1017/S1355617713001483.
12. Dove A, Shang Y, Xu W, Grande G, Laukka EJ, Fratiglioni L, Marseglia A. The impact of diabetes on cognitive impairment and its progression to dementia. Alzheimers Dement. 2021;17:1769–78. https://doi.org/10.1002/alz.12482.
13. Xue M, Xu W, Ou Y-N, Cao X-P, Tan M-S, Tan L, Yu J-T. Diabetes mellitus and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 144 prospective studies. Ageing Res Rev. 2019;55:100944. https://doi.org/10.1016/j.arr.2019.100944.

14. Zhang T, Shaw M, Cherbuin N. Association between type 2 diabetes Mellitus and Brain Atrophy: a Meta-analysis. Diabetes Metab J. 2022;46:781–802. https://doi.org/10.4093/dmj.2021.0189.
15. van Sloten TT, Sedaghat S, Carnethon MR, Launer LJ, Stehouwer CDA. Cerebral microvascular complications of type 2 diabetes: stroke, cognitive dysfunction, and depression. Lancet Diabetes Endocrinol. 2020;8:325–36. https://doi.org/10.1016/S2213-8587(19)30405-X.
16. Moran C, Phan TG, Chen J, Blizzard L, Beare R, Venn A, et al. Brain atrophy in type 2 diabetes: regional distribution and influence on cognition. Diabetes Care. 2013;36:4036–42. https://doi.org/10.2337/dc13-0143.
17. Viigimaa M, Sachinidis A, Toumpourleka M, Koutsampasopoulos K, Alliksoo S, Titma T. Macrovascular complications of type 2 diabetes Mellitus. Curr Vasc Pharmacol. 2020;18:110–6. https://doi.org/10.2174/1570161117666190405165151.
18. Johansen MC, Langton-Frost N, Gottesman RF. The role of Cardiovascular Disease in Cognitive Impairment. Curr Geri Rep. 2020;9:1–9. https://doi.org/10.1007/s13670-020-00309-7.
19. Gupta M, Pandey S, Rumman M, Singh B, Mahdi AA. Molecular mechanisms underlying hyperglycemia associated cognitive decline. IBRO Neurosci Rep. 2023;14:57–63. https://doi.org/10.1016/j.ibneur.2022.12.006.
20. Panico F, Sagliano L, Magliacano A, Santangelo G, Trojano L. The relationship between cognitive reserve and cognition in healthy adults: a systematic review. Curr Psychol. 2022. https://doi.org/10.1007/s12144-022-03523-y.
21. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol. 2012;11:1006–12. https://doi.org/10.1016/S1474-4422(12)70191-6.
22. Yasuno F, Minami H, Hattori H. Interaction effect of Alzheimer's disease pathology and education, occupation, and socioeconomic status as a proxy for cognitive reserve on cognitive performance: in vivo positron emission tomography study. Psychogeriatrics. 2020;20:585–93. https://doi.org/10.1111/psyg.12552.
23. Peña-González P, Mondragón-Maya A, Silva-Pereyra J, Roa-Rojas P. Cognitive Reserve and executive functions in adults with type 2 diabetes. J Diabetes Res. 2020;2020:7941543. https://doi.org/10.1155/2020/7941543.
24. Mathisen J, Jensen AKG, Andersen I, Andersen GS, Hvidtfeldt UA, Rod NH. Education and incident type 2 diabetes: quantifying the impact of differential exposure and susceptibility to being overweight or obese. Diabetologia. 2020;63:1764–74. https://doi.org/10.1007/s00125-020-05150-3.
25. Whitaker SM, Bowie JV, McCleary R, Gaskin DJ, LaVeist TA, Thorpe RJ. The Association between Educational Attainment and Diabetes among men in the United States. Am J Mens Health. 2014;8:349–56. https://doi.org/10.1177/1557988313520034.
26. Agardh E, Allebeck P, Hallqvist J, Moradi T, Sidorchuk A. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. Int J Epidemiol. 2011;40:804–18. https://doi.org/10.1093/ije/dyr029.
27. Steele CJ, Schöttker B, Marshall AH, Kouvonen A, O'Doherty MG, Mons U, et al. Education achievement and type 2 diabetes-what mediates the relationship in older adults? Data from the ESTHER study: a population-based cohort study. BMJ Open. 2017;7:e013569. https://doi.org/10.1136/bmjopen-2016-013569.
28. Zhang J, Chen Z, Pärna K, van Zon SKR, Snieder H, Thio CHL. Mediators of the association between educational attainment and type 2 diabetes mellitus: a two-step multivariable mendelian randomisation study. Diabetologia. 2022. https://doi.org/10.1007/s00125-022-05705-6.
29. Sommer I, Griebler U, Mahlknecht P, Thaler K, Bouskill K, Gartlehner G, Mendis S. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. BMC Public Health. 2015;15:914. https://doi.org/10.1186/s12889-015-2227-y.
30. Letellier N, Ilango SD, Mortamais M, Tzourio C, Gabelle A, Empana J-P, et al. Socioeconomic inequalities in dementia risk among a French population-based cohort: quantifying the role of cardiovascular health and vascular events. Eur J Epidemiol. 2021;36:1015–23. https://doi.org/10.1007/s10654-021-00788-8.
31. Liu C, Ma Y, Hofman A, Waziry R, Koton S, Pike JR, et al. Educational Attainment and Dementia: mediation by Mid-life Vascular Risk factors. Ann Neurol. 2023;94:13–26. https://doi.org/10.1002/ana.26647.
32. Ihle A, Oris M, Fagot D, van der Chicherio C, Sauter J, Kliegel M. Associations of educational attainment and cognitive level of job with old age verbal ability and processing speed: the mediating role of chronic diseases. Appl Neuropsychol Adult. 2018;25:356–62. https://doi.org/10.1080/23279095.2017.1306525.
33. Ihle A, Mons U, Perna L, Oris M, Fagot D, Gabriel R, Kliegel M. The relation of obesity to performance in verbal abilities, Processing Speed, and cognitive flexibility in Old Age: the role of Cognitive Reserve. Dement Geriatr Cogn Disord. 2016;42:117–26. https://doi.org/10.1159/000448916.
34. Yaffe K, Falvey C, Hamilton N, Schwartz AV, Simonsick EM, Satterfield S, et al. Diabetes, glucose control, and 9-year cognitive decline among older adults without dementia. Arch Neurol. 2012;69:1170–5. https://doi.org/10.1001/archneurol.2012.1117.
35. Zheng B, Su B, Price G, Tzoulaki I, Ahmadi-Abhari S, Middleton L, Glycemic, Control. Diabetic complications, and risk of dementia in patients with diabetes: results from a large U.K. Cohort Study. Diabetes Care. 2021;44:1556–63. https://doi.org/10.2337/dc20-2850.
36. Kowall B, Rathmann W. Combined effects of Diabetes and Education on decline of cognitive performance in the older Population: the Survey of Health, Ageing, and Retirement in Europe. Gerontology. 2023;69:172–80. https://doi.org/10.1159/000524571.
37. Nakahori N, Sekine M, Yamada M, Tatsuse T, Kido H, Suzuki M. A pathway from low socioeconomic status to dementia in Japan: results from the Toyama dementia survey. BMC Geriatr. 2018. https://doi.org/10.1186/s12877-018-0791-6.
38. Scholtens S, Smidt N, Swertz MA, Bakker SJL, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol. 2015;44:1172–80. https://doi.org/10.1093/ije/dyu229.
39. Lim YY, Jaeger J, Harrington K, Ashwood T, Ellis KA, Stöffler A, et al. Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian imaging, biomarkers, and lifestyle-rate of change substudy (AIBL-ROCS). Arch Clin Neuropsychol. 2013;28:320–30. https://doi.org/10.1093/arclin/act021.
40. Maruff P, Lim YY, Darby D, Ellis KA, Pietrzak RH, Snyder PJ, et al. Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer's disease. BMC Psychol. 2013;1:30. https://doi.org/10.1186/2050-7283-1-30.
41. Aretz B, Janssen F, Vonk JM, Heneka MT, Boezen HM, Doblhammer G. Long-term exposure to fine particulate matter, lung function and cognitive performance: a prospective Dutch cohort study on the underlying routes. Environ Res. 2021;201:111533. https://doi.org/10.1016/j.envres.2021.111533.
42. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods. 2010;15:309–34. https://doi.org/10.1037/a0020761.
43. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation : R Package for Causal Mediation Analysis. J Stat Soft. 2014. https://doi.org/10.18637/jss.v059.i05.
44. Imai K, Keele L, Yamamoto T, Identification. Inference and sensitivity analysis for Causal Mediation effects. Statist Sci. 2010. https://doi.org/10.1214/10-STS321.
45. Sobel. Direct and Indirect effects in Linear Structural equation models. Sociol Methods Res. 1987;16:155–76. https://doi.org/10.1177/0049124187016001006.
46. Houle J, Lauzier-Jobin F, Beaulieu M-D, Meunier S, Coulombe S, Côté J, et al. Socioeconomic status and glycemic control in adult patients with type 2 diabetes: a mediation analysis. BMJ Open Diabetes Res Care. 2016;4:e000184. https://doi.org/10.1136/bmjdrc-2015-000184.
47. Zeng Y, Wu J, Han Y, Chen F, Chen L, Yang S, Fang Y. Educational disparities in the associations between self-monitoring of blood glucose and glycemic control in type 2 diabetes patients in Xiamen, China. J Diabetes. 2018;10:715–23. https://doi.org/10.1111/1753-0407.12651.
48. Mansur RB, Lee Y, Zhou AJ, Carmona NE, Cha DS, Rosenblat JD, et al. Determinants of cognitive function in individuals with type 2 diabetes mellitus: a meta-analysis. Ann Clin Psychiatry. 2018;30:38–50.
49. Garber AJ, Abrahamson MJ, Barzilay JI, Blonde L, Bloomgarden ZT, Bush MA, CONSENSUS STATEMENT BY THE AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY ON THE COMPREHENSIVE TYPE 2 DIABETES MANAGEMENT ALGORITHM–2016 EXECUTIVE SUMMARY, et al. Endocr Pract. 2016;22:84–113. https://doi.org/10.4158/EP151126.CS.
50. Karter AJ, Stevens MR, Brown AF, Duru OK, Gregg EW, Gary TL, et al. Educational disparities in health behaviors among patients with diabetes: the translating Research Into Action for Diabetes (TRIAD) Study. BMC Public Health. 2007;7:308. https://doi.org/10.1186/1471-2458-7-308.
51. Chen P, Callisaya M, Wills K, Greenaway T, Winzenberg T. Cognition, educational attainment and diabetes distress predict poor health literacy in diabetes: a cross-sectional analysis of the SHELLED study. PLoS ONE. 2022;17:e0267265. https://doi.org/10.1371/journal.pone.0267265.

52. Shrivastava SR, Shrivastava PS, Ramasamy J. Role of self-care in management of diabetes mellitus. J Diabetes Metab Disord. 2013;12:14. https://doi.org/10.1186/2251-6581-12-14.
53. Liao Y-S, Tsai W-C, Chiu L-T, Kung P-T. Educational attainment affects the diagnostic time in type 2 diabetes mellitus and the mortality risk of those enrolled in the diabetes pay-for-performance program. Health Policy. 2023;138:104917. https://doi.org/10.1016/j.healthpol.2023.104917.
54. Yang J, Xu H, Li J, Zhao Y, Guan S, Fu Y, et al. The association between undiagnosed diabetes and cognitive function: findings from the China health and retirement longitudinal study. BMC Endocr Disord. 2022;22:151. https://doi.org/10.1186/s12902-022-01055-x.
55. Guerrero-Berroa E, Ravona-Springer R, Heymann A, Schmeidler J, Levy A, Leroith D, Beeri MS. Haptoglobin genotype modulates the relationships of glycaemic control with cognitive function in elderly individuals with type 2 diabetes. Diabetologia. 2015;58:736–44. https://doi.org/10.1007/s00125-014-3487-2.
56. Baumgaertel J, Haussmann R, Gruschwitz A, Werner A, Osterrath A, Lange J, et al. Education and genetic risk modulate hippocampal structure in Alzheimer's Disease. Aging Dis. 2016;7:553–60. https://doi.org/10.14336/AD.2016.0305.
57. Lee M, Hughes TM, George KM, Griswold ME, Sedaghat S, Simino J, Lutsey PL. Education and Cardiovascular Health as Effect modifiers of APOE ε4 on dementia: the atherosclerosis risk in communities Study. J Gerontol Biol Sci Med Sci. 2022;77:1199–207. https://doi.org/10.1093/gerona/glab299.
58. Kaplan GA, Turrell G, Lynch JW, Everson SA, Helkala EL, Salonen JT. Childhood socioeconomic position and cognitive function in adulthood. Int J Epidemiol. 2001;30:256–63. https://doi.org/10.1093/ije/30.2.256.
59. Bertola L, Benseñor IM, Barreto SM, Giatti L, Moreno AB, Viana MC, et al. Early life socioeconomic status predicts cognition regardless of education level. Eur J Neurol. 2021;28:3972–8. https://doi.org/10.1111/ene.15042.
60. International Diabetes Federation. IDF Diabetes Atlas. 2013.

Publisher's note

9.3. Study 3

Age and Ageing 2022; **51:** 1–9
https://doi.org/10.1093/ageing/afab231
Published electronically 7 December 2021

RESEARCH PAPER

Diabetes duration and the risk of dementia: a cohort study based on German health claims data

Constantin Reinke[1], Nikolaus Buchmann[2], Anne Fink[3], Christina Tegeler[2,4], Ilja Demuth[2,5], Gabriele Doblhammer[1,3]

[1]Institute for Sociology and Demography, University of Rostock, 18057 Rostock, Germany
[2]Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Endocrinology and Metabolic Diseases (including Division of Lipid Metabolism), Biology of Aging working group, 13353 Berlin, Germany
[3]German Center for Neurodegenerative Diseases, 53127 Bonn, Germany
[4]MSB Medical School Berlin, Department of Psychology, 14197 Berlin, Germany
[5]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, BCRT - Berlin Institute of Health Center for Regenerative Therapies, Berlin, Germany

Address correspondence to: Constantin Reinke, Institute for Sociology and Demography, University of Rostock, Ulmenstr. 69, 18057 Rostock, Germany. Tel: +49 (0)381 – 498 4399; Fax: +49 (0) 381 498 - 4395; Email: constantin.reinke@uni-rostock.de

Abstract

Objective: Diabetes is a risk factor for dementia but little is known about the impact of diabetes duration on the risk of dementia. We investigated the effect of type 2 diabetes duration on the risk of dementia.
Design: Prospective cohort study using health claims data representative for the older German population. The data contain information about diagnoses and medical prescriptions from the in- and outpatient sector.
Methods: We performed piecewise exponential models with a linear and a quadratic term for time since first type 2 diabetes diagnosis to predict the dementia risk in a sample of 13,761 subjects (2,558 dementia cases) older than 65 years. We controlled for severity of diabetes using the Adopted Diabetes Complications Severity Index.
Results: We found a U-shaped dementia risk over time. After type 2 diabetes diagnosis the dementia risk decreased (26% after 1 year) and reached a minimum at 4.75 years, followed by an increase through the end of follow-up. The pattern was consistent over different treatment groups, with the strongest U-shape for insulin treatment and for those with diabetes complications at the time of diabetes diagnosis.
Conclusions: We identified a non-linear association of type 2 diabetes duration and the risk of dementia. Physicians should closely monitor cognitive function in diabetic patients beyond the first few years after diagnosis, because the later increase in dementia occurred in all treatment groups.

Keywords: Type 2 diabetes, dementia, risk factors, health claims data, older people

Key Points

- The dementia risk over time since first type 2 diabetes diagnosis was U-shaped.
- The U-shaped pattern was consistent over different diabetes treatment groups.
- Less severe cases with no diabetes complications at baseline showed a stronger U-shape than more severe cases.

Introduction

The prevalence of dementia and diabetes has increased in recent decades [1]. The WHO forecasts an increase in the all-cause dementia prevalence to 152 million by 2050 [2].

Type 2 diabetes (T2D) is a well-known risk factor for dementia [3–8]. The prevalence of T2D was estimated to be 6.28% (462 million people) worldwide, and 22% among people aged 70 years and older [9]. Even pre-diabetes has been shown to be a risk factor for all-cause dementia, Alzheimer's disease (AD) and vascular dementia (VD). Changes of diabetes-related biochemical indicators such as fasting plasma insulin were associated with an increased dementia risk [8]. While cardiovascular risk factors are the main drivers for an increased dementia risk in midlife, diabetes seems to be the strongest predictor among the cardiovascular risk factors in later-life [10].

There are various mechanisms by which T2D and dementia are connected [11]. First, there are atherosclerotic consequences of T2D that favour VD. T2D is strongly associated with micro- and macrovascular diseases. This association is essentially determined by diabetes duration [12, 13].

However, little is known about the association of diabetes duration and the risk of dementia.

Differences in the mean age of dementia onset with respect to diabetes duration attenuated in older age groups [14]. Other studies found an increased risk of cognitive decline in prevalent diabetes rather than in incident diabetes [15] or with longer duration [16]. Studies emphasising the role of age at diabetes onset also provided evidence of increasing dementia risk with longer diabetes duration [17, 18].

In addition to duration and earlier onset of diabetes, the risk of mild cognitive impairment (MCI) increases with the severity of diabetes [19], and diabetes is associated with a higher risk for the progression from MCI to dementia [20]. Baseline HemoglobinA_{1c} (HbA_{1c}) was negatively associated with cognition, supporting the assumption that severity of T2D play a crucial role in the development of dementia [21]. Severe hypoglycaemia also promotes the occurrence of dementia [22]. Peripheral insulin resistance appears to be linked to cerebral insulin resistance, and lower glucose utilisation in the brain may drive ad development [22, 23], and T2D severity and progression are associated with the risk for dementia [24].

Due to dementia subtypes and the role T2D can play in the development of VD and ad, the influence of different treatments for T2D on dementia is complex. It was recently reported that distinct treatments of T2D result in different incident rates of dementia, with the lowest risk in oral antidiabetic medications (ADM), followed by diabetes patients without ADM and those dependent on insulin [7, 25]. However, some studies suggest a potential benefit from diabetes treatment to the risk of AD [26] and all-cause dementia [27, 28]. There is evidence for a reduced dementia risk from metformin [27–29].

The aim of this study was to explore the association between diabetes duration and dementia incidence, taking severity and treatment form of T2D into account.

We hypothesised that the dementia risk increases with longer T2D duration independent of T2D severity. This duration effect should be moderated by the different diabetes treatment strategies.

Methods

Study design and sample

A random sample of 250,000 individuals was drawn in 2004 by the largest German health insurance, the 'Allgemeine Ortskrankenkasse' (AOK). The sample is representative of the German population aged 65 and older, as measured in terms of mortality (Supplementary Figure S1). The claims data were anonymized (by the data provider) and we did not have access to the primary data, so no ethical review or patient consent was required. The sample included insured people born before 1955 with a follow-up to 2015 and contained information about demographic data, diagnosis from in- and outpatient sector based on International Classification of Diseases (ICD-10, [30]), as well as all medical prescriptions according to the Anatomical Therapeutic Chemical Classification (ATC) System [31]. The data were structured on a quarterly basis. We included people with incident T2D diagnosis who were born before 1940 and excluded people with dementia diagnoses or T2D diagnoses before 2006, or people with any type 1 diabetes diagnosis during the follow-up (excluded in this specific order). Our sample consisted of 13,761 insured people older than 64 years with newly diagnosed T2D and at least one quarter of follow-up. We observed 2,558 dementia cases through the end of 2014, 2,845 people died within this period and 8,544 reached the end of follow-up (Supplementary Figure S2), 107 people dropped out of the data for other reasons (e.g. change of insurance company). Data from the year 2015 were used for validation only.

Diabetes and Dementia

T2D and dementia were identified by ICD-10 codes. To reduce the problem of false positive diagnoses, we used internal validation strategies (Supplementary Text S1).

Diabetes treatment

We defined three groups of ADM, coded by ATC-codes, as well as a group without any ADM prescription. Insulin users received prescriptions with ATC-code A10A. The group of non-insulin ADM had prescriptions with ATC-code A10B. We assigned concurrent prescriptions of insulin and non-insulin ADM to the group of mixed ADM.

Covariates

To approximate the severity of diabetes, we implemented the Adopted Diabetes Complications Severity Index (aDCSI) [32–34]. The aDCSI score ranged from 0 to 13. For our analysis we categorised the aDCSI into five levels (0, 1, 2, 3 and 4+).

Further covariates were age, sex, hypertensive diseases, depression, cerebrovascular diseases, ischemic heart diseases, atrial fibrillation and flutter, obesity and disorders of lipoprotein metabolism (ICD-10 codes: Supplementary Text S2).

Statistical analysis

We assessed the risk of dementia depending on the time since first T2D diagnosis. We measured dementia incidence from the quarter of first T2D diagnosis through the last quarter of 2014. In the multivariable analysis we performed piecewise exponential regression models, which permitted us to explicitly model the baseline hazard over the analysis time. We split the baseline hazard into quarters and defined T2D duration as a second-degree polynomial in terms of time since T2D diagnosis (d) and a quadratic term (d^2) (Table 2: Model 1). More details on methods and formulas are provided in the online supplement (Supplementary Text S3).

Treatment strategies were included as a time-varying predictor, thus, individuals were able to change the treatment groups. To avoid biases we followed Hernan et al. [35, 36] (Supplementary Text S4).

The diabetes complications (aDCSI) was measured as a time-varying variable, and a time dummy controlled for a structural changes in the billing system for physicians [37] (Supplementary Text S3).

To examine how the duration effect of T2D differs between treatment strategies of T2D, we included interaction effects between the treatment group (tg_i) and the baseline hazard function (model 4). Formulas are provided in the online supplement (Supplementary Text S3).

In a sensitivity analysis, we performed models separately for age groups (65–84 years and older than 84 years; models 2 and 3, and 5 and 6), and for the diabetes severity at the time of T2D incidence in terms of less severe cases (aDCSI = 0) vs. more severe (aDCSI>0).

We performed Cox regression models to statistically confirm that the effects of treatment strategies and T2D severity on dementia incidence were independent of our model strategy (results upon request).

All analyses were performed using Stata 16.0 (College Station, TX).

Results

Descriptive results

The analysis sample comprised 57,613 person-years. The mean follow-up time per subject was 4.18 years, and the mean age at first T2D diagnosis 76.9 years (sd = 5.8). The mean of the aDCSI at the time of T2D incidence was 2.34 (sd = 1.46) with 2,278 people without any diabetes complications.

The incidence of dementia decreased after the first year, remained nearly constant for the next 3 years (Table 1), and increased thereafter. More severe T2D cases were more likely to receive a new diagnosis of dementia than less severe cases. Dementia incidence was significantly higher in women than in men and increased with age and with almost all the comorbidities considered. Incidence differed significantly among treatment groups, with those treated with insulin having the highest incidence (Table 1).

Model results

Table 2 shows the estimated hazard ratios (HR) of the piecewise exponential models to assess the dementia risk since T2D incidence. In models 1–3 the linear term d of time since T2D incidence revealed a decreasing risk (model 1: HR(d) = 0.92; 95%CI = 0.90–0.93; model 2: HR(d) = 0.92; 95%CI = 0.91–0.94, model 3: HR(d) = 0.90; 95%CI = 0.86–0.93). Furthermore the significant quadratic terms indicated an increase in dementia risk for longer time-spans (model 1: $HR(d^2) = 1.002$; 95%CI = 1.002–1.003; model 2: $HR(d^2) = 1.002$; 95%CI = 1.001–1.002, model 3: $HR(d^2) = 1.003$; 95%CI = 1.001–1.004). This combination resulted in a U-shaped risk pattern for dementia over time. After1year the dementia risk decreased by 26% (predicted $HR(d,d^2) = 0.74$; 95%CI = 0.70–0.78), and reached its minimum 4.75 years after T2D incidence (predicted $HR(d,d^2) = 0.44$; 95%CI = 0.39–0.50, Figure 1). Model 4 included an interaction term between diabetes duration and the treatment groups. We observed a U-shaped risk pattern for diabetic people without any ADM (model reference group: HR(d) = 0.93; 95%CI = 0.91–0.94, and $HR(d^2) = 1.002$; 95%CI = 1.001–1.003), and an increased dementia risk for the group of insulin users within the quarter of T2D incidence (model 4: $HR(d,d^2,tg_1) = 3.61$; 95%CI = 2.30–5.65), with a stronger U-shape thereafter ($HR(d,tg_1) = 0.86$; 95%CI = 0.79–0.94, $HR(d^2,tg_1) = 1.004$; 95%CI = 1.001–1.006). Thus, after 1 year the predicted dementia risk in the group of insulin users was about 60% higher ($HR(d,d^2,tg_1) = 1.61$; 95%CI = 1.21–2.14) than for individuals without ADM. Both other treatment groups did not differ significantly from the group without any ADM. The U-shaped pattern consisted in both age groups as well as in all treatment groups, with the exception of the mixed ADM group (Figure 2). Stratification for diabetes complications at T2D incidence shifted the U-shape upwards for individuals with complications, which we consider to be more severe cases (Figure 3). The initial decrease in the dementia risk in subsequence of incident T2D is stronger for less severe diabetes cases, while the U-shaped pattern for the more severe cases is less pronounced.

Table 1. Characteristics of the study population and dementia incidence rate per 1,000 person-years with 95% confidence intervals

Variable	Person-years	Cases with dementia	Dementia incidence rate per 1,000 person-years		
			Rate	95% Confidence interval	
Time since T2D diagnosis					
Up to 1 year	15,913	816	51.28	47.88	54.92
Up to 2 years	10,672	408	38.23	34.70	42.13
Up to 3 years	8,966	352	39.26	35.37	43.58
Up to 4 years	7,335	270	36.81	32.67	41.47
Up to 5 years	5,713	246	43.06	38.00	48.79
Up to 6 years	4,216	194	46.01	39.97	52.96
Up to 7 years	2,815	153	54.35	46.39	63.69
Up to 8 years	1,574	95	60.35	49.36	73.79
More than 8 years	407	24	58.91	39.48	87.89
Diabetes severity at T2D incidence					
Less severe cases	11,407	323	28.32	25.39	31.58
more severe	46,206	2,235	48.37	46.41	50.42
Sex					
Man	22,568	883	39.13	36.63	41.79
Woman	35,044	1,675	47.80	45.56	50.14
Age group at T2D incidence					
65–84	53,500	2,054	38.39	36.77	40.09
85+	4,112	504	122.56	112.31	133.74
Treatment groups					
No ADM	40,772	1,864	45.72	43.69	47.84
Insulin	1,071	83	77.47	62.47	96.07
Non-insulin ADM	15,044	577	38.35	35.35	41.61
Mixed ADM	725	34	46.88	33.50	65.61
Hypertensive diseases					
No	3,129	126	40.27	33.82	47.95
Yes	54,484	2,432	44.64	42.90	46.45
Depression					
No	39,047	1,444	36.98	35.12	38.94
Yes	18,566	1,114	60.00	56.58	63.63
Cerebrovascular diseases					
No	38,850	1,192	30.68	28.99	32.47
Yes	18,763	1,366	72.80	69.04	76.77
Ischemic heart diseases					
No	26,453	953	36.03	33.81	38.39
Yes	31,160	1,605	51.51	49.05	54.09
Atrial fibrillation and flutter					
No	43,528	1,615	37.10	35.34	38.96
Yes	14,085	943	66.95	62.81	71.36
Obesity					
No	35,852	1,742	48.59	46.36	50.93
Yes	21,761	816	37.50	35.01	40.16
Disorders of lipoprotein metabolism					
No	17,761	880	49.55	46.38	52.93
Yes	39,852	1,678	42.11	40.14	44.17
Total	**57,613**	**2,558**	**44.40**	**42.71**	**46.15**

Discussion

This study provides evidence for a U-shaped association between T2D duration and the dementia risk for individuals older than 64 years with newly diagnosed T2D. The identified U-shaped pattern was independent of the severity of diabetes. After the initial T2D diagnosis, the dementia risk continued to decrease for 5 years, followed by an increase thereafter. However, even over longer durations the risk did not surpass levels observed immediately after T2D diagnosis. This pattern held true for both age groups. Interaction between the duration of T2D and different treatment strategies did not change the U-pattern for three of four treatment strategies and disclosed the strongest U-shape for insulin treatment. Only the group with mixed treatment strategy did not reveal a clear U-pattern. To consider different stages of T2D progression at the time of first diagnosis, we used stratified models by diabetes-severity. We found a stronger U-shape for less severe cases at baseline than for more severe cases.

Our results support previous findings from Chiu et al. [24], who identified diabetes progression significantly

Table 2. Results of regression models, hazard ratios and 95% confidence intervals, risk of dementia dependent on the duration of diabetes

Variable	Hazard Ratio (95% CI)	Hazard Ratio (95% CI)	Hazard Ratio (95% CI)
	Model 1 (all)	**Model 2 (65–84 years)**	**Model 3 (85+)**
Time since T2D (d)	0.918***	0.924***	0.899***
	(0.904–0.932)	(0.909–0.940)	(0.864–0.935)
Time since T2D² (d^2)	1.002***	1.002***	1.003***
	(1.002–1.003)	(1.001–1.002)	(1.001–1.004)
	Model 4 (all)	**Model 5 (65–84 years)**	**Model 6 (85+ years)**
Time since T2D (d)	0.926***	0.933***	0.908***
	(0.910–0.943)	(0.914–0.952)	(0.870–0.948)
Time since T2D² (d^2)	1.002***	1.002***	1.003***
	(1.001–1.003)	(1.001–1.002)	(1.001–1.004)
Treatment (tg_i) (ref. no ADM tg_0)			
Insulin (tg_1)	3.609**	3.819***	4.347***
	(2.305–5.651)	(2.289–6.372)	(1.634–11.565)
Non-insulin ADM (tg_2)	1.168	1.164	1.119
	(0.926–1.471)	(0.893–1.519)	(0.674–1.858)
Mixed ADM (tg_3)	2.186*	2.030	2.937
	(1.000–4.775)	(0.811–5.077)	(0.657–13.122)
Treatment (tg_i) × Time since T2D (d)			
Insulin (d, tg_1)	0.862***	0.883***	0.660***
	(0.794–0.937)	(0.807–0.966)	(0.495–0.879)
Non-insulin ADM (d, tg_2)	0.981	0.979	1.002
	(0.945–1.018)	(0.940–1.019)	(0.901–1.115)
Mixed ADM (d, tg_3)	0.966	0.981	0.883
	(0.849–1.100)	(0.849–1.134)	(0.631–1.237)
Treatment (tg_i) × Time since T2D² (d^2)			
Insulin (d^2, tg_1)	1.004***	1.003**	1.014**
	(1.001–1.006)	(1.00–1.006)	(1.003–1.024)
Non-insulin ADM (d^2, tg_2)	1.000	1.001	0.999
	(0.999–1.002)	(0.999–1.002)	(0.995–1.004)
Mixed ADM (d^2, tg_3)	0.999	0.999	1.002
	(0.995–1.004)	(0.994–1.003)	(0.990–1.014)

95% CI: 95% Confidence interval. *P value <0.10. **P value <0.05. ***P value <0.01. All models controlled for: age, sex, comorbidity and aDCSI. Models 1, 2 and 3 explore the total duration effect. Models 4, 5 and 6 explore the duration effect by treatment groups.

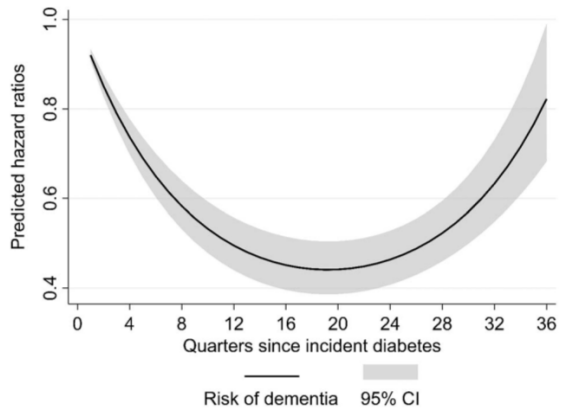


Figure 1. Predicted risk of dementia over time since incident T2D diagnosis. CI = Confidence intervals. Source: AOK data 2004–2015, authors' calculations.

associated with dementia risk. Contrary to these findings we also identified the T2D duration as a risk factor, independent of the diabetes severity and unrelated to the progression of T2D. Our results contrast with other findings from Wu et al. [16], who did not indicate a higher risk for cognitive decline in incident diabetes patients.

In general, a worsening of the cerebrovascular blood supply can promote loss of cognitive function. Mild chronic inflammatory constellations in the context of insulin resistance favour this process [38]. The U-shape may be caused by several factors. In Germany, exposure to medical services increases in advanced age, particularly among men [39]. This can generally contribute to an increased detection rate of dementia during check-up or other examinations, so it may be a monitoring effect that need not be limited to diabetes. Moreover, the German Diabetes Association recommends close-meshed tests for dementia in T2D subjects [40], which may explain the initial high dementia risk. Pre-diabetes

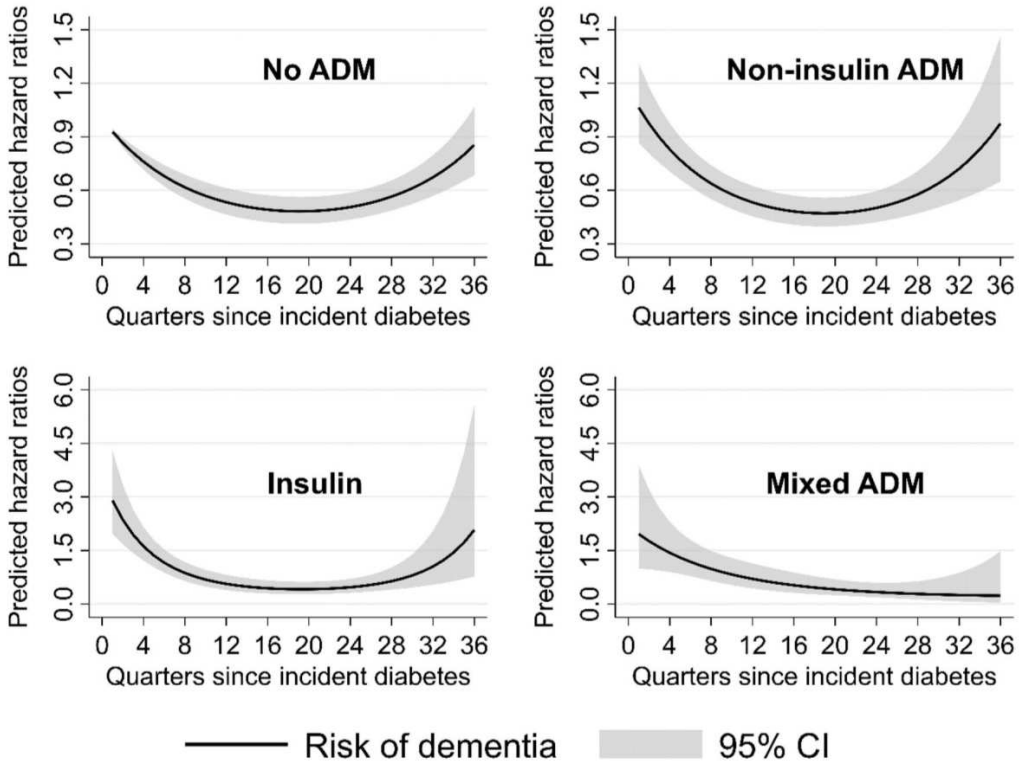


Figure 2. Predicted risk of dementia over time since incident T2D diagnosis for treatment groups. CI = Confidence intervals. Note different y-axis scales. Source: AOK data 2004–2015, authors' calculations.

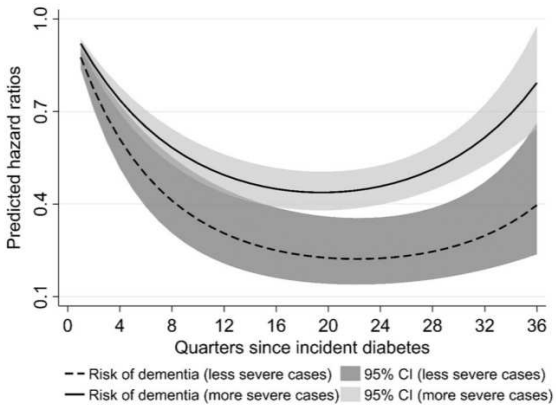


Figure 3. Predicted risk of dementia over time since T2D by diabetes severity at T2D incidence. Source: AOK data 2004–2015, authors' calculations.

is associated with an increased risk of impaired cognitive function [8], which could contribute to the high dementia prevalence at T2D diagnosis. However, when glycaemic control has not yet been established, a diagnosis of dementia cannot be made without restrictions. A newly diagnosed T2D is usually accompanied by start of a therapy, which consists of medical treatment and/or changes in lifestyle. One aspect for the initial decrease in dementia diagnosis might be that a formal diagnosis of dementia in subjects with poorly controlled T2D will be delayed until good glycaemic control has been established. This should, however, only affect a small number of patients and a diagnosis of dementia after drug or lifestyle interventions is usually made before the increase observed here. The start of diabetes therapy may reduce or delay cognitive decline [41] in the prodromal phase [26] and might be partially responsible for a delayed development/diagnosis of dementia. Notably, patient compliance with diabetes self-management decreases over time [42] and reduces this effect. Also, loss of beta cell function continues contributing to the need for more medication [43]. Expanding treatment complexity, not limited to diabetes, may play a role here. [44–46]. It is also conceivable that there is decreased motivation to adhere to treatment recommendations due to an abatement of perceived treatment efficacy. It appears that if the diagnosis of dementia is not made within the first year after T2D diagnosis, this happens hesitantly for a longer time. Diagnosis of dementia can be challenging as early dementia symptoms may not be apparent during time-limited physician–patient interactions. Furthermore, there might be a low awareness for cognitive decline in younger subjects. Accordingly, early dementia symptoms often remain undetected by physicians and patients resulting

in delayed diagnosis. Cognitive decline may worsen self-management of T2D therapy, even in early and undetected stages of dementia. This is a vulnerable phase, as cognitive changes may be aggravated by incorrect self-management of T2D as a consequence of hypo- or hyperglycaemia [47, 48].

Over time micro- and macrovascular complications develop in T2D and may outweigh the therapy effect. This suspicion is supported by the weaker shaped U-form for the severe diabetes cases at baseline (Figure 3). The incidence of dementia diagnoses starting five or more years after T2D diagnosis could thus express the long-term consequences of T2D on cognition.

The lack of a clear U-shaped pattern in the mixed group may be caused by heterogeneous composition in terms of T2D severity and treatment, and needs further investigation.

Strengths and limitations

Unlike several smaller previous studies [14, 16, 19] and similar to Chiu et al. [24] we analysed a large longitudinal sample with up to 9 years of follow-up.

As these health claims data are not based on interviews and recruiting, recall and selection bias as well as panel attrition can be ruled out. Patients were included regardless of their cognitive and functional status, which is particularly important for the oldest ones who are living in nursing homes.

We used a validated measure for the diabetes severity (aDCSI) as a time-varying covariate. This allowed us to consider the severity and progression of T2D.

Our study is not without limitations. We observed the time of diagnoses rather than the time of onset. T2D and dementia are both slowly progressive diseases and the lag time between onset and first diagnoses can vary considerably [49, 50], as can the level of diabetes complications [51].

Diagnoses are made not only through physician services, which can lead to underdiagnosis, but also through a higher rate of dementia detection through the use of health care services not limited to diabetes. These diagnoses tools are neither standardised nor always specific, which may lead to inaccurately encoded diagnoses. To reduce this problem, we applied established internal validation strategies for diabetes and dementia. In case of dementia more than 50% of diagnoses were coded as 'unspecific' (ICD-F03), hence we cannot distinguish between dementia subtypes, although T2D might affect VD in a different way than ad [11].

Information about drug use is restricted to prescriptions filled in a pharmacy and we have no information on whether medication was taken. A significant proportion of T2D patients might actually be affected by latent autoimmune diabetes in adults (LADA), a type of diabetes that could not be distinguished in the current study. The prevalence of LADA is estimated between 2 and 14% [52], and it cannot be ruled out that this affected the results. Finally, the data do not contain information about socio-economic background, lifestyle factors and health behaviour. While changes in lifestyle are usually part of diabetes treatment strategies, we do not know whether these were followed. We cannot draw conclusions about glycaemic control as we have no information on blood glucose levels.

Conclusion

The main finding of the current study is the U-shaped risk of dementia over time from the time of diabetes diagnosis. Possible explanations for the initially high incidence of dementia include a better screening in T2D patients, a consequence of a deteriorated metabolic situation even before T2D diagnosis and a higher exposure to medical services in this population. The following decrease in dementia incidence 2–5 years after diagnosis of T2D could be due to lifestyle and drug intervention but might also depend on higher awareness for dementia in T2D subjects. It might be assumed that after a longer exposure to T2D and diabetic complications the incidence of dementia raises over time starting approximately 5–8 years after the initial diagnosis of T2D. The aspects discussed above certainly cannot fully explain the U-shaped dementia risk over time in patients following a diabetes diagnosis, and further research is needed to obtain a more comprehensive picture and derive practical implications for the prevention and treatment of cognitive impairment in T2D patients. Our data suggest that physicians should be encouraged to continue close monitoring of the development of cognitive function in diabetic patients even if diabetes was diagnosed more than 2 years ago. In this context, treatment outcome and adherence should also be considered. Future research should investigate whether primary prevention, and detection and treatment of T2D might be beneficial not only regarding the development of T2D, but also regarding the development of dementia.

Supplementary Data: Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

Declaration of Conflicts of Interest: None.

Declaration of Sources of Funding: C.R. and G.D. were part of the project 'Multi-state, multi-time, multi-level analysis of health-related demographic events: Statistical aspects and applications', which was funded by the German Research Foundation (Funding program 'Research Grants' https://gepris.dfg.de/ project number 386913674). The funders did not have any role in study design, collection, analysis and interpretation of data, in writing of the report or in the decision to submit the article for publication.

Acknowledgment: We are grateful to the Scientific Research Institute of the AOK, WIdO, for providing the data. The Scientific Research Institute of the AOK (WIdO) imposes strict rules on sharing health claims data as these are classified according to ethical restrictions due to privacy concerns. Anonymised data are available to researchers and institutions upon request. In order to request access to the health claims

data of the AOK, please contact the WIdO directly (http://www.wido.de/, mail: wido@wido.bv.aok.de).

References

1. Nichols E, Szoeke CEI, Vollset SE *et al.* Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the global burden of disease study 2016. The Lancet Neurology 2019; 18: 88–106.
2. World Health Organization. Dementia. Key facts. World Health Organization, 2021. https://www.who.int/en/news-room/fact-sheets/detail/dementia (5 March 2021, date last accessed).
3. Gudala K, Bansal D, Schifano F, Bhansali A. Diabetes mellitus and risk of dementia: a meta-analysis of prospective observational studies. J Diabetes Investig 2013; 4: 640–50.
4. Chatterjee S, Peters SAE, Woodward M *et al.* Type 2 diabetes as a risk factor for dementia in women compared with men: a pooled analysis of 2.3 million people comprising more than 100,000 cases of dementia. Diabetes Care 2016; 39: 300–7.
5. Cheng G, Huang C, Deng H, Wang H. Diabetes as a risk factor for dementia and mild cognitive impairment: a meta-analysis of longitudinal studies. Intern Med J 2012; 42: 484–91.
6. Haroon NN, Austin PC, Shah BR, Wu J, Gill SS, Booth GL. Risk of dementia in seniors with newly diagnosed diabetes: a population-based study. Diabetes Care 2015; 38: 1868–75.
7. Buchmann N, Fink A, Tegeler C, Demuth I, Doblhammer G, Steinhagen-Thiessen E. Different treatment forms of type II diabetes and the risk of dementia in German health claims data. Acta Diabetol 2019; 56: 995–1003.
8. Xue M, Xu W, Ou Y-N *et al.* Diabetes mellitus and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 144 prospective studies. Ageing Res Rev 2019; 55: 100944.
9. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. J Epidemiol Glob Health 2020; 10: 107–11.
10. Kloppenborg RP, van den Berg E, Kappelle LJ, Biessels GJ. Diabetes and other vascular risk factors for dementia: which factor matters most? A systematic review. Eur J Pharmacol 2008; 585: 97–108.
11. Biessels GJ, Despa F. Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications. Nat Rev Endocrinol 2018; 14: 591–604.
12. Zoungas S, Woodward M, Li Q *et al.* Impact of age, age at diagnosis and duration of diabetes on the risk of macrovascular and microvascular complications and death in type 2 diabetes. Diabetologia 2014; 57: 2465–74.
13. Nanayakkara N, Ranasinha S, Gadowski A *et al.* Age, age at diagnosis and diabetes duration are all associated with vascular complications in type 2 diabetes. J Diabetes Complications 2018; 32: 279–90.
14. Zilkens RR, Davis WA, Spilsbury K, Semmens JB, Bruce DG. Earlier age of dementia onset and shorter survival times in dementia patients with diabetes. Am J Epidemiol 2013; 177: 1246–54.
15. Croxson SC, Jagger C. Diabetes and cognitive impairment: a community-based study of elderly subjects. Age Ageing 1995; 24: 421–4.
16. Wu Q, Tchetgen Tchetgen EJ, Osypuk T *et al.* Estimating the cognitive effects of prevalent diabetes, recent onset diabetes, and the duration of diabetes among older adults. Dement Geriatr Cogn Disord 2015; 39: 239–49.
17. Barbiellini Amidei C, Fayosse A, Dumurgier J *et al.* Association between age at diabetes onset and subsequent risk of dementia. JAMA 2021; 325: 1640–9.
18. Xu W, Qiu C, Gatz M, Pedersen NL, Johansson B, Fratiglioni L. Mid- and late-life diabetes in relation to the risk of dementia: a population-based twin study. Diabetes 2009; 58: 71–7.
19. Roberts RO, Geda YE, Knopman DS *et al.* Association of duration and severity of diabetes mellitus with mild cognitive impairment. Arch Neurol 2008; 65: 1066–73.
20. Velayudhan L, Poppe M, Archer N, Proitsi P, Brown RG, Lovestone S. Risk of developing dementia in people with diabetes and mild cognitive impairment. Br J Psychiatry 2010; 196: 36–40.
21. Cukierman-Yaffe T, Gerstein HC, Williamson JD *et al.* Relationship between baseline glycemic control and cognitive function in individuals with type 2 diabetes and other cardiovascular risk factors: the action to control cardiovascular risk in diabetes-memory in diabetes (ACCORD-MIND) trial. Diabetes Care 2009; 32: 221–6.
22. Lee AK, Rawlings AM, Lee CJ *et al.* Severe hypoglycaemia, mild cognitive impairment, dementia and brain volumes in older adults with type 2 diabetes: the atherosclerosis risk in communities (ARIC) cohort study. Diabetologia 2018; 61: 1956–65.
23. Koekkoek PS, Kappelle LJ, van den Berg E, Rutten GEHM, Biessels GJ. Cognitive function in patients with diabetes mellitus: guidance for daily care. The Lancet Neurology 2015; 14: 329–40.
24. Chiu W-C, Ho W-C, Liao D-L *et al.* Progress of diabetic severity and risk of dementia. J Clin Endocrinol Metab 2015; 100: 2899–908.
25. McMillan JM, Mele BS, Hogan DB, Leung AA. Impact of pharmacological treatment of diabetes mellitus on dementia risk: systematic review and meta-analysis. BMJ Open Diabetes Res Care 2018; 6: e000563.
26. McIntosh EC, Nation DA. Importance of treatment status in links between type 2 diabetes and Alzheimer's disease. Diabetes Care 2019; 42: 972–9.
27. Chin-Hsiao T. Metformin and the risk of dementia in type 2 diabetes patients. Aging Dis 2019; 10: 37–48.
28. Orkaby AR, Cho K, Cormack J, Gagnon DR, Driver JA. Metformin vs sulfonylurea use and risk of dementia in US veterans aged ≥65 years with diabetes. Neurology 2017; 89: 1877–85.
29. Campbell JM, Stephenson MD, de Courten B, Chapman I, Bellman SM, Aromataris E. Metformin use associated with reduced risk of dementia in patients with diabetes: a systematic review and meta-analysis. J Alzheimers Dis 2018; 65: 1225–36.
30. World Health Organization. International statistical classification of diseases and related health problems. Geneva: World Health Organization, 2004.
31. World Health Organization Centre for Drug Statistics Methodology. Guidelines for ATC classification and DDD assignment 2018. In: Guidelines for ATC classification and DDD assignment 2018. Oslo, Norway: WHO Collaborating Centre for Drug Statistics Methodology, 2017.

32. Young BA, Lin E, von Korff M *et al.* Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. Am J Manag Care 2008; 14: 15–23.

33. Chang H-Y, Weiner JP, Richards TM, Bleich SN, Segal JB. Validating the adapted diabetes complications severity index in claims data. Am J Manag Care 2012; 18: 721–6.

34. Wicke FS, Glushan A, Schubert I *et al.* Performance of the adapted diabetes complications severity index translated to ICD-10. Am J Manag Care 2019; 25: e45–9.

35. Emilsson L, García-Albéniz X, Logan RW, Caniglia EC, Kalager M, Hernán MA. Examining bias in studies of statin treatment and survival in patients with cancer. JAMA Oncol 2017; 4: 63–70.

36. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016; 79: 70–5.

37. Schulz M, Goffrier B, Stillfried DV *et al.* Entwicklung der vertragsärztlichen Leistungsdichte bei Diagnostik und Therapie der Demenz – Update 2010 bis 2014. Berlin, Germany: Zentralinstitut für die kassenärztliche Versorgung in Deutschland, 2017.

38. Ninomiya T. Diabetes mellitus and dementia. Curr Diab Rep 2014; 14: 487.

39. Robert Koch-Institut. Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2010". Berlin: Robert-Koch-Inst, 2012.

40. Kulzer B, Albus C, Herpertz S *et al.* Psychosoziales und diabetes (Teil 1). Diabetologie und Stoffwechsel 2013; 8: 198–242.

41. Yaffe K, Falvey C, Hamilton N *et al.* Diabetes, glucose control, and 9-year cognitive decline among older adults without dementia. Arch Neurol 2012; 69: 1170–5.

42. Shrivastava SR, Shrivastava PS, Ramasamy J. Role of self-care in management of diabetes mellitus. J Diabetes Metab Disord 2013; 12: 14.

43. Elder DA, Hornung LN, Khoury JC, D'Alessio DA. β-Cell function over time in adolescents with new type 2 diabetes and obese adolescents without diabetes. J Adolesc Health 2017; 61: 703–8.

44. Coleman CI, Limone B, Sobieraj DM *et al.* Dosing frequency and medication adherence in chronic disease. J Manag Care Pharm 2012; 18: 527–39.

45. Polonsky WH, Henry RR. Poor medication adherence in type 2 diabetes: recognizing the scope of the problem and its key contributors. Patient Prefer Adherence 2016; 10: 1299–307.

46. Chapman CD, Schiöth HB, Grillo CA, Benedict C. Intranasal insulin in Alzheimer's disease: food for thought. Neuropharmacology 2018; 136: 196–201.

47. Hopkins R, Shaver K, Weinstock RS. Management of Adults with diabetes and cognitive problems. Diabetes Spectr 2016; 29: 224–37.

48. Puttanna A, Padinjakara NK. Management of diabetes and dementia. Br J Diabetes 2017; 17: 93.

49. Harris MI, Klein R, Welborn TA, Knuiman MW. Onset of NIDDM occurs at least 4-7 yr before clinical diagnosis. Diabetes Care 1992; 15: 815–9.

50. Karr JE, Graham RB, Hofer SM, Muniz-Terrera G. When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death. Psychol Aging 2018; 33: 195–218.

51. Porta M, Curletto G, Cipullo D *et al.* Estimating the delay between onset and diagnosis of type 2 diabetes from the time course of retinopathy prevalence. Diabetes Care 2014; 37: 1668–74.

52. Hernández M, Mauricio D. Latent autoimmune diabetes in adults: a review of clinically relevant issues. Adv Exp Med Biol 2021; 1307: 29–41.

Received 21 May 2021; editorial decision 24 September 2021

9.4. Study 4

Acta Psychiatrica Scandinavica

Acta Psychiatr Scand 2020: 142: 402–412
DOI: 10.1111/acps.13220

ACTA PSYCHIATRICA SCANDINAVICA

The temporal association between incident late-life depression and incident dementia

Heser K, Fink A, Reinke C, Wagner M, Doblhammer G. The temporal association between incident late-life depression and incident dementia.

Objective: There is an established association between depression and subsequent dementia. The present study examined temporal associations between incident late-life depression and subsequent dementia, also considering age and sex.
Methods: We used longitudinal health claims data from the largest German health insurance provider ('Allgemeine Ortskrankenkasse') considering up to 9 follow-up years in piecewise exponential models. ICD-10 codes were used to define incident depression and dementia in individuals ≥65 years (n = 97 110).
Results: Incident depression was associated with a higher risk of subsequent dementia (incidence rate ratios (IRR) adjusted for age and sex: IRR = 1.58, 95% CI = 1.51–1.64). The strongest association was found for the shortest interval of 1 quarter (IRR = 2.04, 95% CI = 1.88–2.21), with significant associations up to an interval of roughly 3 years. The association was more pronounced and lasted for more quarters in the younger portion of this study group (ages from 65–74: IRR = 2.00, 95% CI = 1.83–2.18; 75–84: IRR = 1.64, 95% CI = 1.55–1.73; ≥85: IRR = 1.19, 95% CI = 1.08–1.31). It was stronger among men than women (men: IRR = 1.98, 95% CI = 1.84–2.14; women: IRR = 1.44, 95% CI = 1.37–1.51) with no sex-specific temporal association.
Conclusion: This large claims data study confirmed that incident late-life depression is associated with a higher risk of dementia within the 3 years following diagnosis. Hence, incident late-life depression should prompt further cognitive examinations and referrals to specialists. This might apply especially to younger seniors and men.

K. Heser[1], A. Fink[2], C. Reinke[3], M. Wagner[1,2], G. Doblhammer[2,3]

[1]Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, North Rhine-Westphalia, 53127, Germany, [2]German Center for Neurodegenerative Diseases (DZNE), Bonn, North Rhine-Westphalia, Germany and [3]University of Rostock, Rostock, Mecklenburg-West Pomerania, Germany

Key words: depression; dementia; epidemiology

Kathrin Heser, Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, North Rhine-Westphalia, Germany. E-mail: Kathrin.Heser@ukbonn.de

Accepted for publication July 20, 2020

Significant outcomes

- An association between incident late-life depression and a higher risk of subsequent dementia was found.
- The strongest association between incident late-life depression and subsequent dementia risk was found for the shortest interval between both diagnoses.
- A stronger association between incident late-life depression and subsequent dementia risk was found in younger old individuals and in men.

Limitations

- Although registry-based data have several strengths, it is also limited to general sociodemographic sample characteristics, for example, without information on neuropsychological performance.
- False diagnoses of incident depression or dementia might explain the short temporal association between both disorders, although we do not consider this likely due to our validation strategy of diagnoses.
- Our results might not be generalizable to members of other public or private health insurance groups.

Introduction

In aging societies, dementia is an increasingly prevalent condition incurring high societal costs and placing additional burdens on patients and their relatives (1). Depression can be both a risk factor for, and a prodromal feature of dementia (2) in different individuals and samples. Recent meta-analyses have found that, on average, late-life depression or clinically relevant depressive symptoms were associated with a twofold higher risk of incident dementia (3, 4). Importantly, this association seemed to be time-dependent and may be moderated by age and sex. Further elucidation of these moderators could yield clinically useful information, in that incident depression in some patient groups may come along with a higher risk of dementia within close temporal proximity, prompting clinical attention.

Previous research has shown that depressive symptoms at both mid- and late-life were associated with a higher risk of dementia (5), whereas others found a higher risk of subsequent dementia specifically for participants with late-onset depression and elevated late-life depressive symptoms (6, 7). Several studies have suggested a close temporal proximity between late-life or late-onset depression and subsequent dementia (8–12) and intervals of up to 5 years might be particularly relevant (13–15). These results point to a prodromal association between depression and dementia, in which depression can be an early sign of the yet latent and preclinical phase of the dementia process. However, some studies (16, 17) have indicated there may be associations for considerably longer periods between both diagnoses. The patients' age at diagnosis may also affect the temporal association between incident late-life depression and dementia. Only a few studies have addressed this issue so far, with inconsistent results. In a population study with subjects over 75 years of age, Heser et al. (7) found that the risk of subsequent all-cause dementia was higher with increasing age at depression onset. It was insignificant for a recalled depression onset at age 60 and increased to a significant risk with a later onset, for example, at 75 years. Conversely, in a prospective study, Chen et al. (18) found a stronger association between depressive syndromes and dementia in a group of 65- to 74-year-olds with decreasing and non-significant associations for the elder groups. However, these latter results refer to late-life depression in general rather than incident late-life depression. A higher prevalence of depression in women compared with men has often been found (e.g., (19)). It has also been suggested that women have a higher risk of dementia and Alzheimer's disease (AD), but results are less consistent than those for depression and might be driven by women's higher life-expectancy and by geographical differences (20). Sex-specific risk factors of dementia, especially for AD, have recently attracted more interest (20, 21). It is still an open issue whether the association between depression and subsequent dementia differs between men and women. While some studies reported that depression is more strongly associated with a higher risk of dementia in men (13,22–24), others found an association only in women (25, 26), or did not find any sex differences (18, 27).

Aims of the study

In sum, while temporal and demographic variables might be interesting moderators of the depression-dementia link, the existing literature does not yet allow researchers and clinicians to draw solid conclusions. Sample sizes have often been small, rendering stratified analyses prone to inconclusive findings. In addition, different age ranges and differences regarding the definition of depression (e.g., symptoms versus diagnosis and incident vs. prevalent depression) limit the integration and interpretation of the data which had been collected. We sought to examine this issue by leveraging a large dataset containing clinician-coded depression and dementia diagnoses. We used registry-based longitudinal health claims data, approximately representative of the German population aged 65 and above. With regard to the existing literature on the prodromal association between depression and dementia, we hypothesized that associations would be stronger for shorter intervals between incident depression and subsequent dementia. In order to better characterize patient groups with incident depression at risk of dementia, we explored whether this temporal association regarding time intervals of quarters between both diagnoses of incident depression and subsequent dementia may be moderated by age of depression incidence and sex. Our results might assist diagnostic and treatment decisions of clinicians regarding different age and sex strata.

Material and methods

Study design and sample

We analyzed health claims data from the largest German health insurance provider, the 'Allgemeine Ortskrankenkasse' (AOK). A random sample of 250 000 insured persons born before 1955 was drawn by the data provider in 2004 with a

follow-up through 2015. The data provide information about sex, year, and month of birth and death, in- and outpatient diagnoses coded according to the 10th revision of the International Classification of Diseases (ICD-10), the specialty of physicians, and all medical prescriptions. All medical information was available on a quarterly basis. We included persons who were born before 1940 and excluded persons with dementia or depression diagnoses in 2004 or 2005. Our analysis sample consisted of 97 110 insured persons older than 64 years and included 20 779 dementia cases through the end of 2014 (Fig. 1). This study involved anonymized claims data and fell outside the scope of the Declaration of Helsinki and did not require ethical review.

Depression diagnosis

We used the following ICD-10 codes for the identification of valid depression diagnoses: depressive episode (F32); recurrent depressive disorder (F33); and dysthymia (F34.1). However, F33 and F34.1 were used only for validation of F32 diagnoses. The validation of depression diagnoses was based on a strategy described in (28). A diagnosis was assumed to be valid if one of the following conditions was met: an inpatient discharge or secondary diagnosis; two outpatient diagnoses in different quarters (within four quarters time); or two outpatient diagnoses by separate physicians in the same quarter. A diagnosis was also assumed to be valid if the person died in the same quarter of the initial depression diagnosis. The severity of depression is classified into broader general categories (mild, moderate, severe, or unknown) which were based on ICD-10 codes (e.g., F32.1). Cases with a depression diagnosis in 2004 or 2005 were excluded from our analyses, as we could not know the time of the first depression incidence. All cases with a valid depression diagnosis after 2005 were considered to be incident cases. We try to approximate incident cases of depression by applying a two-year washout-period, although it is possible that there had been incident cases of depression diagnosed before 2004. Persons remain in the exposed group from the time of first diagnosis onwards; in a sensitivity analysis, we started from the second diagnosis. Remission from depression was not coded.

Dementia diagnosis

Dementia diagnoses were defined by one of the following ICD-10 codes: dementia in AD (F00), without dementia in AD with early onset (F00.0); vascular dementia (F01); dementia in other diseases classified elsewhere (F02); unspecified

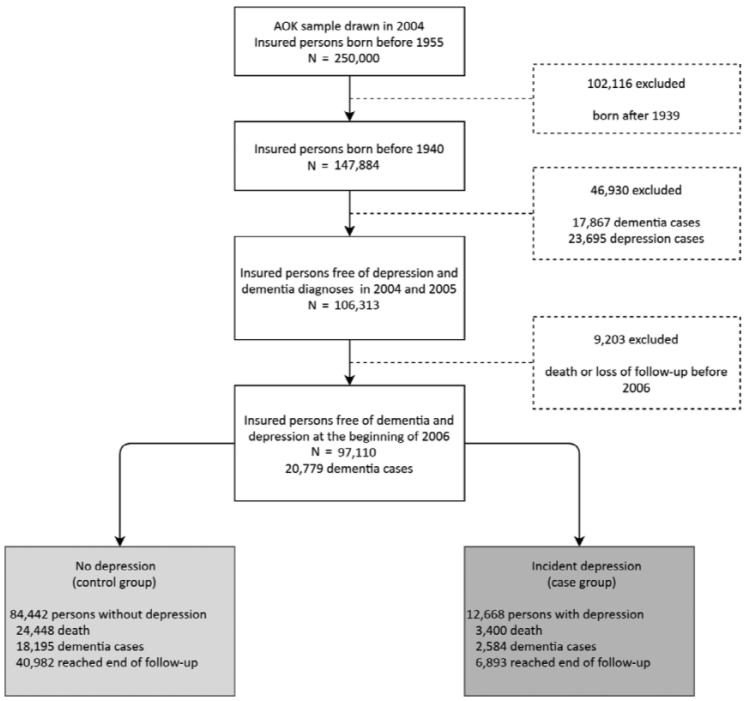


Fig. 1. Inclusion of insured subjects in the analyses.

dementia (F03); delirium superimposed on dementia (F05.1); and AD (G30), without AD with early onset (G30.0). We applied a two-stage validation strategy to overcome the problem of false-positive diagnoses (29). In the first stage, 'verified' outpatient diagnoses as well as inpatient discharge or secondary diagnoses were identified. In the second stage, the diagnoses were confirmed by a simultaneous in- and outpatient diagnosis, by two diagnoses from two different types of physicians in the same quarter, or by co-occurrence over the study period plus the year 2015, which was only used for validation. A diagnosis was also considered to be valid if a person died in the same quarter as the initial dementia diagnosis. If cases validated by co-occurrence over the study period, the time of incident dementia was set to the time of first diagnosis. In a sensitivity analysis, we used time of the second diagnosis. As for depression, all persons with a dementia diagnosis in 2004 or 2005 were excluded. All persons with a valid dementia diagnosis after 2005 were considered to be incident cases.

Covariates

Covariates were age, sex, and each patient's history of comorbidities, including: mental and behavioral disorders due to psychoactive substance use (F10-F19); schizophrenia, schizotypal, and delusional disorders (F20-F29); hypertensive diseases (I10-I15); diabetes mellitus without type I (E11-E14); cerebrovascular diseases (I60-I69); ischemic heart diseases (I20-I25); injuries to the head (S00-S09); neurotic, stress-related, and somatoform disorders (F40); other anxiety disorders (F41); reaction to severe stress and adjustment disorders (F43); bipolar affective disorder (F31); and atrial fibrillation and flutter (I48). All covariates, with the exception of sex, were considered to be time-varying variables with a value of 1 at time of the first diagnosis in the data.

Statistical analysis

For analyzing the temporal association of dementia and depression, we assessed the risk of dementia depending on the depression status. We calculated the dementia incidence rate and performed piecewise constant exponential models. Dementia incidence was measured in the 9-year period from the first quarter of 2006 through the last quarter of 2014. Following Hernán et al. (2016) and Emilsson et al. (2018), we considered depression to be a time-varying variable (30, 31). Therefore, we doubled all eligible observations (measured in person-times) and assigned each copy either to the control or to the case group (Fig. 1). For both groups, we defined a time zero of analysis time. For person-times without a depression diagnosis, analysis time started at the beginning of 2006. In the case group, time zero was defined as the time of incident depression diagnosis (middle of the relevant quarter). That means that a person could start in the control group and be censored at time of the first depression diagnosis. This person could then switch to the case group, where his or her analysis time was set to zero again. All persons (in both the control and case groups) were censored at the time of death or at the end of follow-up (end of 2014). Analysis time was measured in months and was split into a maximum of 36-time intervals (quarters) in order to perform piecewise exponential models, which allow a varying baseline hazard. Using the Akaike information criteria to evaluate different functions for the temporal specification, we determined that the natural logarithm of the analysis time best represented our data. The time of dementia diagnosis was set to the middle of the relevant quarter. The time of death was set to the middle of the month of death. We included cases with concurrent onset of depression and dementia in the control group. All analyses were performed using Stata 16.0.

Model 1 was adjusted for age and sex, model 2 additionally for comorbidities. Further, we performed sex- and age-specific models to assess the dementia risk for men and women separately (models 3 and 4) and for the three age groups 65–74 years, 75–84 years, and 85+ years (models 5 to 7). The assignment to one of these age groups was based on the age at time zero for each group. To compare the dementia risk of the control and case group over time, models 8 to 14 replicated models 1 to 7 and estimated additionally the time-related dementia risks by including interaction effects between depression (yes ? case group/no ? control group) and the analysis time. All models included a time-dummy for the fourth quarter in 2013 and the first quarter of 2014 to consider structural changes in the billing system for physicians ('Chronikerzuschlag'; (32)). This change led to a temporary increase of diagnoses of chronic diseases which do not reflect the epidemiological development in Germany.

Results

Descriptive results

Our analyzed sample comprised 97,110 persons with 636 322 person years. The mean follow-up time per subject was 5.82 years. The mean age at

start of observation was 74.7 years (SD = 6.6) for subjects without depression and 78.1 years (SD = 6.1) for subjects at the time of incident depression. Table 1 shows characteristics of the study population and the rate of dementia incidence. There were 12 668 subjects with incident depression (41 384 person years) and 2584 of them subsequently developed dementia (Fig. 1 and Table 1). Incident depression almost doubled the incidence of dementia, and the increase rose with depression severity. Dementia incidence rate was slightly higher in women than in men, and it increased with age and all the comorbidities considered, with the exception of neurotic, stress-related, and somatoform disorders (Table 1).

Table 1. Characteristics of the study population and dementia incidence rate per 1000 person years with 95% confidence intervals

			Dementia incidence rate per 1000 person years		
Variable	Person years	Cases with dementia	Rate	95% Confidence interval	
Depression					
No	594 938	18 195	30.58	30.14	31.03
Yes	41 384	2584	62.44	60.08	64.89
Severity of depression					
Unknown	24 861	1809	72.77	69.49	76.20
Mild	3991	347	86.94	78.26	96.59
Moderate	7612	751	98.66	91.85	105.97
Severe	5030	567	112.73	103.82	122.40
Sex					
Male	262 952	7618	28.97	28.33	29.63
Female	373 371	13 161	35.25	34.65	35.86
Age group at start of observation					
65-74	386 055	6156	15.95	15.55	16.35
75-84	216 258	10 722	49.58	48.65	50.53
85+	34 010	3901	114.70	111.16	118.36
Mental & behavioral disorders due to psychoactive substance use					
No	617 183	19 462	31.53	31.09	31.98
Yes	19 139	1317	68.81	65.19	72.63
Schizophrenia					
No	628 241	19 770	31.47	31.03	31.91
Yes	8081	1009	124.86	117.39	132.81
Hypertensive diseases					
No	110 456	2044	18.51	17.72	19.32
Yes	525 866	18 735	35.63	35.12	36.14
Diabetes without type I					
No	402 467	10 933	27.16	26.66	27.68
Yes	233 855	9846	42.10	41.28	42.94
Cerebrovascular diseases					
No	480 089	10 534	21.94	21.53	22.36
Yes	156 234	10 245	65.57	64.32	66.86
Ischemic heart diseases					
No	366 959	9138	24.90	24.40	25.42
Yes	269 363	11 641	43.22	42.44	44.01
Injuries to the head					
No	564 148	15 985	28.33	27.90	28.78
Yes	72 174	4794	66.42	64.57	68.33
Neurotic, stress-related disorders					
No	623 767	20 371	32.66	32.21	33.11
Yes	12 555	408	32.50	29.49	35.81
Other anxiety disorders					
No	598 295	19 068	31.87	31.42	32.33
Yes	38 027	1711	44.99	42.91	47.18
Reaction to severe stress, and adjustment disorders					
No	587 021	18 833	32.08	31.63	32.54
Yes	49 302	1946	39.47	37.76	41.26
Bipolar affective disorder					
No	635 464	20 701	32.58	32.14	33.02
Yes	859	78	90.85	72.77	113.42
Atrial fibrillation and flutter					
No	530 131	14 194	26.77	26.34	27.22
Yes	106 191	6585	62.01	60.53	63.53
Total	636 322	20 779	32.65	32.21	33.10

Source: AOK data 2004–2015, own calculations.

Model results

Table 2 shows the estimated incidence rate ratios (IRR) of the piecewise exponential models to assess the risk of dementia depending on an incident depression diagnosis. Incident depression increased the subsequent risk of dementia by a factor of 1.58 ($P < 0.01$, model 1). Controlling for comorbidities attenuated some of this effect (IRR = 1.11, $P < 0.01$, model 2). Men with incident depression had a 1.98-fold ($P < 0.01$, model 3), and women with incident depression had a 1.44-fold ($P < 0.01$, model 4) higher risk of subsequent dementia. In the age group 65–74, an incident depression diagnosis increased the risk of dementia by a factor of 2 ($P < 0.01$, model 5), persons aged 75–84 years had a 1.64-fold ($P < 0.01$, model 6) higher risk, and persons aged 85 years or above had a 1.19-fold ($P < 0.01$, model 7) higher risk. We confirmed the effect of depression on dementia using Cox models in a sensitivity analysis. However, since Cox regression does not permit us to estimate the temporal association of depression and dementia, we continue with piecewise exponential models.

Temporal association of depression and dementia

Figures 2 and 3 show the predicted incidence rates of the total interaction effect concerning the dementia risk of depressed and non-depressed persons for each time interval of the analysis time (adjusted for age and sex for all, adjusted for age in sex-specific models, adjusted for age and sex in age-specific models). In the first quarter after an incident depression diagnosis, the risk of dementia was 2.04-fold (calculation in supplementary material, see Table S1 and formula below) compared to the first quarter after study entry of persons without depression (Table 2, model 8; Fig. 2, all). This excess risk decreased with increasing follow-up time. After 11 quarters (about 3 years), depressed and non-depressed persons did not differ significantly regarding their dementia risk. After 20

Table 2. Results of regression models representing risk of subsequent dementia depending on incident depression diagnosis

	Model 1† (no comorb.)	Model 2‡ (with comorb.)	Model 3§ (male)	Model 4§ (female)	Model 5† (65–74 years)	Model 6† (75–84 years)	Model 7† (85 + years)
Variable	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
Depression	1.58 (1.51–1.64)	1.11 (1.07–1.17)	1.98 (1.84–2.14)	1.44 (1.37–1.51)	2.00 (1.83–2.18)	1.64 (1.55–1.73)	1.19 (1.08–1.31)
	Model 8† (no comorb.)	Model 9‡ (with comorb.)	Model 10§ (male)	Model 11§ (female)	Model 12† (65–74 years)	Model 13† (75–84 years)	Model 14† (85 + years)
Depression	2.04 (1.88–2.21)	1.38 (1.27–1.49)	2.61 (2.27–3.01)	1.82 (1.65–2.01)	3.84 (3.17–4.64)	2.30 (2.06–2.56)	1.34 (1.14–1.57)
Ln(Time)	0.84 (0.82–0.86)	0.76 (0.74–0.77)	0.88 (0.84–0.91)	0.83 (0.81–0.85)	0.90 (0.84–0.96)	0.86 (0.83–0.89)	0.83 (0.78–0.88)
Depression ϑ Ln(Time)	0.79 (0.74–0.84)	0.82 (0.77–0.88)	0.77 (0.68–0.86)	0.80 (0.74–0.87)	0.61 (□ 0.53 to 0.69)	0.73 (0.67–0.80)	0.87 (0.75–1.02)

95% CI, 95% Confidence interval; IRR, Incidence Rate Ratio; comorb, comorbidity.
Source: AOK data 2004–2015, own calculations.
All p-values are ≤ 0.05 (with the exception of Depression ϑ Ln(Time) in model 14).
†Controlled for age and sex.
‡Controlled for age, sex, and comorbidities.
§Controlled for age.

quarters, both groups had the same dementia risk. In the first quarter, depressed men had a 2.61-fold higher dementia risk compared with non-depressed men (Fig. 2, male). After 13 quarters, depressed and non-depressed men did not differ significantly. Depressed women started with a 1.82-fold higher dementia risk. Depressed and non-depressed women did not differ significantly regarding their dementia risk after 8 quarters of follow-up (Fig. 2, female).

Persons aged 65–74 years with a diagnosis of depression had a 3.84-fold higher dementia risk compared to non-depressed persons of the same age group (Fig. 3, 65–74). After 9 quarters, the two groups did not differ significantly. Depressed persons aged 75–84 years had a 2.3-fold higher dementia risk and after 8 quarters did not significantly differ from non-depressed persons of this age group (Fig. 3, 75–84). Depressed persons aged 85 years and above had only a 1.34-fold higher

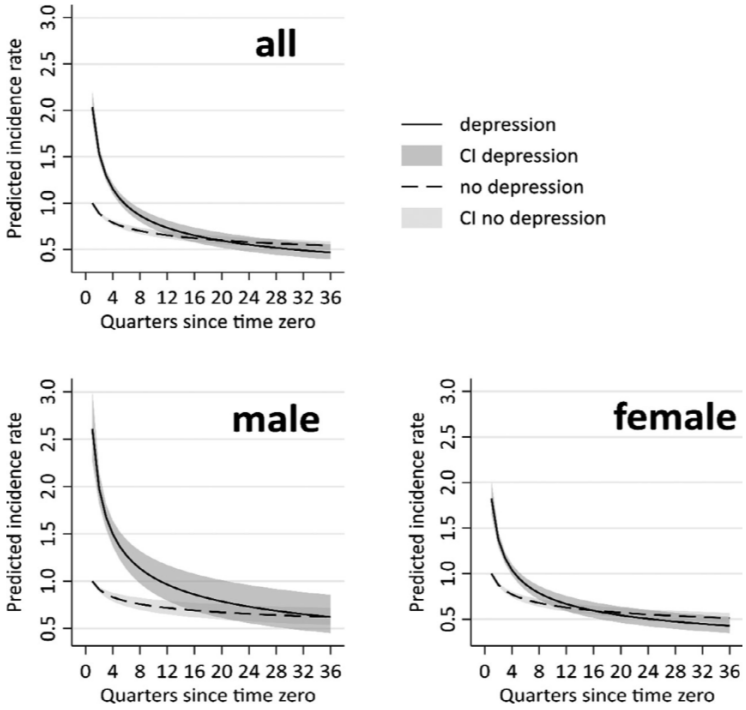


Fig. 2. Predicted incidence rates with 95% confidence intervals for dementia risk of persons with and without depression dependent on time since time zero for the total study population (all), men (male), and women (female). Source: AOK data 2004–2015, own calculations.

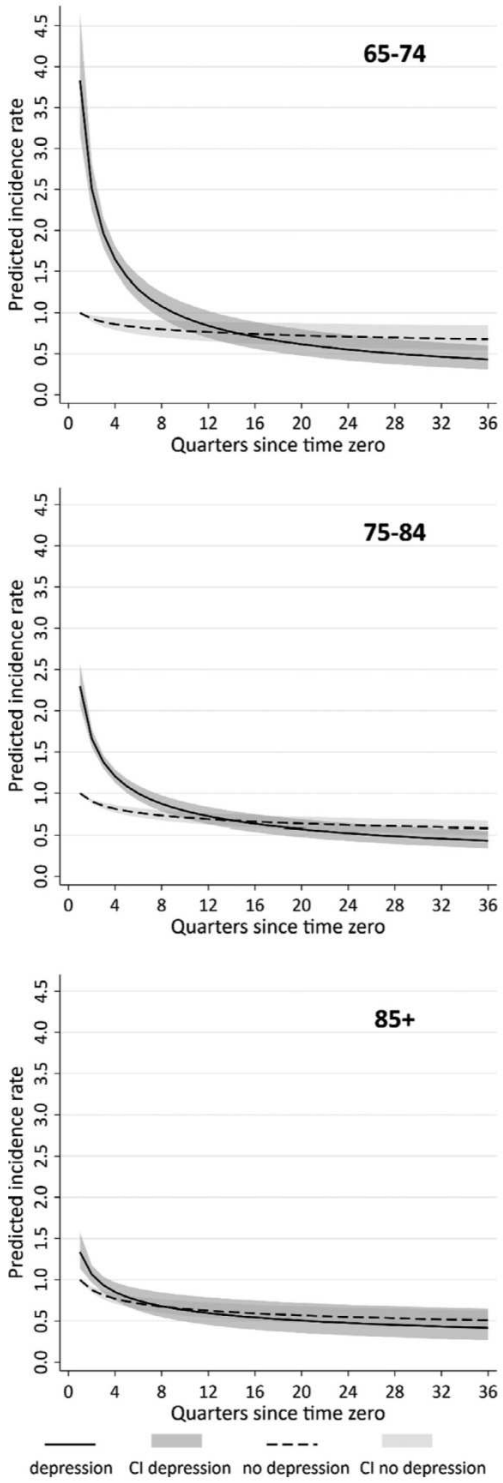


Fig. 3. Predicted incidence rates with 95% confidence intervals for dementia risk of persons with and without depression dependent on time since time zero by age groups (65–74, 75–84, 85+). Source: AOK data 2004–2015, own calculations.

dementia risk and did not significantly differ after 3 quarters (Fig. 3, 85+).

Adjustment for comorbidities attenuates some but not all of the duration effect.

Discussion

Using a large claims data sample from the largest German health insurer, we clearly confirmed that a new diagnosis of depression in old age is associated with a higher risk of incident dementia within the subsequent 3 years, but this association ceases thereafter. This temporal association was particularly strong between ages 65 and 74 and became weaker with advancing age; above age 85 it was only present up to three quarters after the depression diagnosis. To our knowledge, this is the first study that has shown significant differences in the pattern of the temporal association between incident late-life depression and subsequent dementia in different age groups and for both sexes.

The 2.04-fold risk observed for receiving a subsequent dementia diagnosis within the next quarter, which leveled off but remained significant for about 3 years, confirms results from a previous registry-based study: Tapiainen et al. (15) found that the association between hospital-discharge diagnoses of 'depression and other mood disorders' and subsequent risk of AD weakened when the time period between both diagnoses increased from 0 years (OR about 1.7) to 7 years (OR about 1.1). Our data confirm and extend this finding to the general population, including persons with a diagnosis of depression who were not hospitalized. A marked temporal decay of the association between any newly diagnosed mood disorder and any organic disorder was also reported by Plana-Ripoll et al. (33), with strong associations particularly in the first 6 months. Prospective epidemiological studies have shown an increase of depressive symptoms as early as a decade before dementia, reaching a threshold of 'clinically relevant depression' about 4 years before (cf. (12)). Another population-based study found only an association between the increasing depressive symptoms trajectory and subsequent dementia risk compared to the constantly low depressive symptom trajectory, whereas the decreasing, remitting, and constantly high depressive symptoms trajectories were not significantly associated with incident dementia risk over a 10-year follow-up (34). Methodological characteristics of registry-based and epidemiological studies might contribute to the temporal association between depression and dementia. As we used clinical diagnoses, cases of depression and dementia are presumably more

severe or characterized by a later disease stage than those identified by screening scales or diagnostic interviews in epidemiological studies (35, 36). Additionally, not all subjects who score above a depression screener cut-off receive a depression diagnosis in the healthcare setting. However, those diagnosed with depression in health care are probably more likely to be treated and monitored. This results in relatively earlier detection of dementia, for example, when treatment is not effective and additional problems are reported during follow-up visits. It is therefore plausible that registry and health claims data show a rather close temporal coupling of depression and dementia within a time span of up to 3 years. Together with the epidemiological literature, these data support the assumption that late-life depression can be a prodrome of dementia, which also does not exclude consideration of mid-life depression as a risk factor. Early detection and treatment of depression in the elderly, for example, with selective serotonin reuptake inhibitors (SSRIs), might have a positive effect on cognitive deterioration and incident dementia (37), although well-conducted randomized clinical trials are needed for further examination.

In our study, adjusting for a wide range of mental and physical comorbidities attenuated, but did not eliminate, the association. The attenuation is plausible considering interrelationships of both depression and dementia with vascular diseases (2) or other comorbidities, suggesting that common mechanisms may give rise to incident depression and dementia.

Incident depression and subsequent dementia were most strongly associated in the age group of 65 to 74, with steadily decreasing but still significant associations in the 75 to 84 and 85+ year-old groups. This is in line with the results of Chen et al. (18), who found a stronger association between prevalent depressive syndromes and incident dementia in the age group 65–74. We were also able to show that with increasing age the temporal association between depression and dementia was reduced from about 3 years to 3 quarters. Several reasons may contribute to this age-dependency. First, there might be a possible bias due to referral and treatment, which could lead to different likelihoods of being diagnosed and treated in case of depression. For example, depression being diagnosed and treated might be more likely in women or in the younger elderly, leading to the underestimation of the association between depression and dementia in the other groups. In older age groups, incident depression may arise more often in the context of loss experiences, multi-morbidity, and increasing disability compared with younger ages (38). In contrast, younger patients with an incipient neurodegenerative disorder may perceive the increasing cognitive impairment as being particularly abnormal, giving rise to depression, which is diagnosed well before the threshold of dementia is reached (e.g., (39)). Finally, due to the high incidence of dementia in very old subjects, older patients presenting with depressive symptoms may have been screened for cognitive impairment more often, and a dementia diagnosis might have been reached without a delay.

We found that depression was a stronger predictor for subsequent dementia in men, which is in line with prior smaller studies (e.g., (23, 24)). In addition, we were able to show that the weakening of the association over time was similar for both sexes, despite higher male susceptibility. The higher base rate of depression in women might partially explain this finding, as women are affected by depression about twice as often as men (19). Furthermore, there may be sex-dependent reporting biases, that is, a higher reluctance to report depressive symptoms among males (24), so that men who receive a depression diagnosis might be more severely affected. As depression severity is related to dementia risk (e.g., (18)), an effect which can be seen also in our data (Table 1), such reporting biases regarding depressive symptoms may also impact the association of depression diagnosis and dementia.

The temporal effect is additive to the age and sex effects described above, resulting in effects for specific subgroups which might also become useful for clinical use. For example, a new diagnosis of depression in the younger senior group (65–74 years) confers about a 3.8-fold risk of incident dementia in the next quarter, and a roughly 2.6-fold risk for men. Another study also found a stronger short-term temporal proximity between depressive symptoms and dementia in men (13), regarding a follow-up time of 0 to 5 years. Thus, any incident late-life depression in younger seniors and in men might serve as a clinical signal for incipient dementia, triggering low-cost cognitive screening for cognitive impairment. However, as depression was also significantly associated with a higher risk of subsequent dementia in women and in older seniors, these groups should not be dismissed. Our data also suggest that a depression diagnosis made longer than 3 years ago is not related to a higher risk of dementia.

The strengths of our study include the large study population and a prospective observation period of about 9 years. Furthermore, the data contained information extracted from both

outpatient and institutionalized populations, which is important when assessing age-related diseases such as dementia. All of these strengths also favor the generalizability of our results. Due to the nature of claims data, positive sample selection bias, reporting bias, and recall bias are less relevant compared with other study designs. Established strategies to validate depression and dementia diagnoses were applied, and we used a time-varying exposure definition in order to prevent immortal time bias (40).

This study is not without limitations. Firstly, as the association between depression and subsequent dementia was strongest for short time intervals, false diagnoses may explain our findings (i.e., true dementia cases initially misdiagnosed as false-depression cases). However, our validation strategy of diagnoses included a confirmation of the diagnoses by a secondary physician, which reduced the probability of misdiagnosis. Secondly, although we introduced a two-year washout-period to exclude prevalent cases of depression, undetected prevalent cases may still bias our results. Thirdly, individuals who receive one diagnosis may have a higher probability of receiving a second diagnosis due to the fact that they are being treated and are in contact with the healthcare system. Related to this point is that our validation procedure sets the incidence of the disease to the time of the first diagnosis, which may introduce an immortal time bias. In a sensitivity analysis, we used the second diagnosis as the incidence of the disease and found that the risk of dementia in subsequence of depression was even higher than in our original model while the pattern over time remained nearly unchanged (results in supplementary material, see Table S2). Fourthly, medical claims data include only the medical histories of those persons who sought medical attention, which lowers the generalizability of our findings. Furthermore, our results might not be applicable to patients with private insurance or who are covered by a smaller health insurance provider, as we used data from only one public insurer in Germany. Studies have shown that the AOK's proportion of persons with a low socio-economic status is higher on average, which may lead to higher morbidity rates compared to other public insurers and also compared to private insurers, which are not included here (41). Unfortunately, our registry-based data do not contain any information on educational attainment, socio-economic status, or lifestyle variables, such as smoking and physical activity, all of which relate to depression and dementia and therefore might confound results. We tried to capture this bias by including several comorbidities which are in turn associated with a person's socio-economic status, although we have to admit that this correction cannot fully address the lack of confounder variables such as smoking and socio-economic status. In addition, we can assume that the relative associations between depression and dementia also apply to the total population. Patients with early-stage dementia may have been undiagnosed in our study. Finally, the data do not allow us to distinguish between dementia subtypes with high accuracy.

We provide a detailed analysis of the association between depression and subsequent dementia in old age. While most cases of incident depression in the elderly are not followed by dementia in the next decade, a clinically relevant incident depression, as noticed and diagnosed by healthcare professionals in daily routine, can be a prodrome of dementia and signal an increased dementia risk in the near future. Such a diagnosis should prompt diagnostic scrutiny and monitoring for cognitive and functional impairment. This is particularly true for the younger elderly population, where the prevalence of dementia is still low. Whether immediate and successful treatment of late-life depression can reduce the risk of dementia is an important field of future research.

Acknowledgements

We are grateful to the Scientific Research Institute of the AOK, WIdO, for providing the data.

Kathrin Heser and Michael Wagner are part of the study 'Healthy Aging: Gender specific trajectories into latest life' (AgeDifferent.de), which was funded by the Bundesministerium für Bildung und Forschung (Funding program 'Gesund – ein Leben lang', grant numbers 01GL1714A, 01GL1714B, 01GL1714C, 01GL1714D). Constantin Reinke and Gabriele Doblhammer are part of the project 'Multi-state, multi-time, multi-level analysis of health-related demographic events: Statistical aspects and applications', which was funded by the German Research Foundation (Funding program 'Research Grants' https://gepris.dfg.de/ project number 386913674). The funders did not have any role in study design, in the collection, analysis, and interpretation of data, in the writing of the report, nor in the decision to submit the article for publication. Open access funding enabled and organized by Projekt DEAL.

Conflict of interest

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: Kathrin Heser and Michael Wagner received financial support from the German Federal Ministry of Education and Research as part of the study 'Healthy Aging: Gender specific trajectories into latest life' (AgeDifferent.de; funding program 'Gesund – ein Leben lang', grant numbers 01GL1714A, 01GL1714B, 01GL1714C, 01GL1714D); Constantin Reinke and Gabriele Doblhammer received financial support from the German Research Foundation (Funding program 'Research Grants' https://gepris.dfg.de/ project number 386913674); Constantin Reinke received personal fees from Heidelberg Engineering

GmbH outside the submitted work; Gabriele Doblhammer received personal fees from Eli Lilly and Company outside the submitted work; there were no financial relationships with any organization that might have an interest in the submitted work in the previous 3 years; and there were no other relationships or activities that could appear to have influenced the submitted work.

Data availability statement

The Scientific Research Institute of the AOK (WIdO) imposes strict rules on sharing health claims data as these are classified according to ethical restrictions due to privacy concerns. Anonymized data are available to researchers and institutions upon request. In order to request access to the health claims data of the AOK, please contact the WIdO directly (http://www.wido.de/, mail: wido@wido.bv.aok.de).

References

1. Wimo A, Guerchet M, Ali GC et al. The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimers Dement 2017;13:1–7.
2. Byers AL, Yaffe K. Depression and risk of developing dementia. Nat Rev Neurol 2011;7:323–331.
3. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: Systematic review and meta-analysis of community-based cohort studies. Br J Psychiatry 2013;202:329–335.
4. Cherbuin N, Kim S, Anstey KJ. Dementia risk estimates associated with measures of depression: A systematic review and meta-analysis. BMJ Open 2015;5:e008853.
5. Barnes DE, Yaffe K, Byers AL, McCormick M, Schaefer C, Whitmer RA. Midlife vs late-life depressive symptoms and risk of dementia: Differential effects for Alzheimer disease and vascular dementia. Arch Gen Psychiatry 2012;69:493–498.
6. Li G, Wang LY, Shofer JB et al. Temporal relationship between depression and dementia: Findings from a large community-based 15-year follow-up study. Arch Gen Psychiatry 2011;68:970–977.
7. Heser K, Tebarth F, Wiese B et al. Age of major depression onset, depressive symptoms, and risk for subsequent dementia: Results of the German study on Ageing, Cognition, and Dementia in Primary Care Patients (AgeCoDe). Psychol Med 2013;43:1597–1610.
8. Steffens DC, Plassman BL, Helms MJ, Welsh-Bohmer KA, Saunders AM, Breitner JCS. A twin study of late-onset depression and apolipoprotein E e4 as risk factors for Alzheimer's disease. Biol Psychiatry 1997;41:851–856.
9. Andersen K, Lolk A, Kragh-Sørensen P, Petersen NE, Green A. Depression and the risk of Alzheimer disease. Epidemiology 2005;16:233–238.
10. Brommelhoff JA, Gatz M, Johansson B, McArdle JJ, Fratiglioni L, Pedersen NL. Depression as a risk factor or prodromal feature for dementia? Findings in a population-based sample of Swedish twins. Psychol Aging 2009;24:373–384.
11. Karlsson IK, Bennet AM, Ploner A et al. Apolipoprotein E e4 genotype and the temporal relationship between depression and dementia. Neurobiol Aging 2015;36:1751–1756.
12. Singh-Manoux A, Dugravot A, Fournier A et al. Trajectories of depressive symptoms before diagnosis of dementia: a 28-year follow-up study. JAMA Psychiatry 2017;74:712–718.
13. Mirza SS, de Bruijn RFAG, Direk N et al. Depressive symptoms predict incident dementia during short- but not long-term follow-up period. Alzheimers Dement 2014;10:S323–S329.e1.
14. Almeida OP, Hankey GJ, Yeap BB, Golledge J, Flicker L. Depression as a modifiable factor to decrease the risk of dementia. Transl Psychiatry 2017;7:e1117.
15. Tapiainen V, Hartikainen S, Taipale H, Tiihonen J, Tolppanen AM. Hospital-treated mental and behavioral disorders and risk of Alzheimer's disease: A nationwide nested case-control study. Eur Psychiatry 2017;43:92–98.
16. Green RC, Cupples LA, Kurz A et al. Depression as a risk factor for Alzheimer disease: The MIRAGE Study. Arch Neurol 2003;60:753–759.
17. Ownby RL, Crocco E, Acevedo A, John V, Loewenstein D. Depression and risk for Alzheimer disease: systematic review, meta-analysis, and metaregression analysis. Arch Gen Psychiatry 2006;63:530–538.
18. Chen R, Hu Z, Wei L, Qin X, McCracken C, Copeland JR. Severity of depression and risk for subsequent dementia: Cohort studies in China and the UK. Br J Psychiatry 2008;193:373–377.
19. Bromet E, Andrade LH, Hwang I et al. Cross-national epidemiology of DSM-IV major depressive episode. BMC Med 2011;9:90.
20. Mielke MM, Ferretti MT, Iulita MF, Hayden K, Khachaturian AS. Sex and gender in Alzheimer's disease – Does it matter? Alzheimers Dement 2018;14:1101–1103.
21. Ferretti MT, Iulita MF, Cavedo E et al. Sex differences in Alzheimer disease – the gateway to precision medicine. Nat Rev Neurol 2018;14:457–469.
22. da Silva J, Gonçalves-Pereira M, Xavier M, Mukaetova-Ladinska EB. Affective disorders and risk of developing dementia: systematic review. Br J Psychiatry 2013;202:177–186.
23. Fuhrer R, Dufouil C, Dartigues JF. Exploring sex differences in the relationship between depressive symptoms and dementia incidence: Prospective results from the PAQUID Study. J Am Geriatr Soc 2003;51:1055–1063.
24. Dal Forno G, Palermo MT, Donohue JE, Karagiozis H, Zonderman AB, Kawas CH. Depressive symptoms, sex, and risk for Alzheimer's disease. Ann Neurol 2005;57:381–387.
25. Artero S, Ancelin ML, Portet F et al. Risk profiles for mild cognitive impairment and progression to dementia are gender specific. J Neurol Neurosurg Psychiatry 2008;79:979–984.
26. Kim S, Kim MJ, Kim S et al. Gender differences in risk factors for transition from mild cognitive impairment to Alzheimer's disease: a CREDOS study. Compr Psychiatry 2015;62:114–122.
27. Kessing LV, Nilsson FM. Increased risk of developing dementia in patients with major affective disorders compared to patients with other medical illnesses. J Affect Disord 2003;73:261–269.
28. Wagner CJ, Dintsios CM, Metzger FG et al. Longterm persistence and nonrecurrence of depression treatment in Germany: A four-year retrospective follow-up using linked claims data. Int J Methods Psychiatr Res 2018;27:e1607.
29. Taylor DH, Østbye T, Langa KM, Weir D, Plassman BL. The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. J Alzheimers Dis 2009;17:807–815.
30. Emilsson L, García-Albéniz X, Logan RW, Caniglia EC, Kalager M, Hernán MA. Examining bias in studies of statin treatment and survival in patients with cancer. JAMA Oncol 2018;4:63–70.

31. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70–75.
32. Schulz M, Goffrier B, von Stillfried D, Bätzing-Feigenbaum J. Entwicklung der vertragsärztlichen Leistungsdichte im Bereich der Demenzdiagnostik und der ärztlichen nicht-medikamentösen Demenztherapie-Update 2010 bis 2014. Berlin: Zentralinstitut für die kassenärztliche Versorgung in Deutschland (Zi) Versorgungsatlas-Bericht Nr 17/04; 2017.
33. Plana-Ripoll O, Pedersen CB, Holtz Y et al. Exploring comorbidity within mental disorders among a danish national population. JAMA Psychiatry 2019;76:259–270.
34. Mirza SS, Wolters FJ, Swanson SA et al. 10-year trajectories of depressive symptoms and risk of dementia: a population-based study. Lancet Psychiatry 2016;3:628–635.
35. Bock C, Bukh JD, Vinberg M, Gether U, Kessing LV. Validity of the diagnosis of a single depressive episode in a case register. Clin Pract Epidemiol Ment Health 2009;5:4.
36. Rizzuto D, Feldman AL, Karlsson IK, Dahl Aslan AK, Gatz M, Pedersen NL. Detection of dementia cases in two Swedish health registers: a validation study. J Alzheimers Dis 2018;61:1301–1310.
37. Bartels C, Wagner M, Wolfsgruber S, Ehrenreich H, Schneider A. Impact of SSRI therapy on risk of conversion from mild cognitive impairment to Alzheimer's dementia in individuals with previous depression. Am J Psychiatry 2018;175:232–241.
38. Blazer DG. Depression in late life: review and commentary. J Gerontol A Biol Sci Med Sci 2003;58:249–265.
39. Jorm AF. Is depression a risk factor for dementia or cognitive decline? A review. Gerontology 2000;46:219–227.
40. Lévesque LE, Hanley JA, Kezouh AK, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ 2010;340:b5087.
41. Schnee M. Sozioökonomische Strukturen und Morbidität in den gesetzlichen Krankenkassen. In: Böcken J, Braun B, Amhof R, editors. Gesundheitsmonitor 2008: Gesundheitsversorgung und Gestaltungsoptionen aus der Perspektive der Bevölkerung. Verlag Bertelsmann Stiftung; 2008: pp. 88–104.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Parameter estimates (b) of regression models representing risk of subsequent dementia depending on incident depression diagnosis.
Table S2. Results of regression models representing risk of subsequent dementia depending on incident depression diagnosis using the alternative validation.