Exploring factors associated with dementia diagnosis and post-diagnosis outcomes in the Electronic Health Care Record data.

Kamran Khan

Institute of Health Informatics

Thesis for Doctor of Philosophy (Ph.D.)

I, Kamran Khan confirm that the work presented in this thesis is my own. Information derived from other sources has been referenced and given credit.

Abstract

Importance

A better understanding of dementia progression will yield important information for future observational studies and clinical trials which will be essential for the approval of disease-modifying therapy (DMT). The knowledge gained will also inform guidelines for the treatment and management of dementia patients and improve resource allocation.

Objectives

- To comprehensively characterize longitudinal cohort studies and to identify what has been studied with regards to dementia progression after diagnosis in individuals aged ≥ 65, highlight limitations, any unexplored areas and identify opportunities to help inform new observational studies and clinical trials.
- To study dementia patients in the electronic health records data and provide an overview of their characteristics and select research-quality cohort.
- Conduct a time-to-event analysis for incident dementia and its risk factors based on the results of our preliminary analysis of our data.
- Describe and use multi-state modeling approach to study post-dementia outcomes using hospital admission and discharge data.

Methods

- For the review of literature, I searched OVID-MEDLINE for longitudinal studies with human participants from April 2008 till April 2019. Studies measuring outcomes of different domains important in dementia progression (clinical, health system utilization, biomarkers) were included.

- Preliminary analysis was performed using primary care general practice data and hospital episode statistics data and different summary statistics were performed and looked at the comorbidities in dementia cohort in these datasets.
- For the time-to-event analysis we used Cox proportional hazard model to study the factors for incident dementia in diabetic patients. We chose diabetes because it was the most prevalent comorbid conditions in dementia patients in the results of literature review and in the preliminary analysis of our datasets.
- For the post-dementia hospitalisation, institutionalisation, and mortality of dementia patients, I used a multi-state Cox model to study these outcomes and risk factors associated simultaneously in one model.

Findings

I included 100 longitudinal studies comprising >2m individuals in the literature review. Mostly they had a small sample size (57% N<500 participants), short follow-up (66% <= 3 years), and dominance of AD (85% of the total sample in the selected studies was AD and only 9% was of vascular dementia (VaD)). Studies were mainly focused on measuring cognition (69% studies), while functioning and quality of life were less commonly measured (45% and 9% studies respectively). Studies were mainly measuring outcomes at 1 to 3 different time points and the follow-ups were shorter.

The percentage of incident dementia in diabetic patients during the 10-years follow-up was 18.9 cases per 1000-person years. Increasing age, female gender and diabetes duration were associated with higher risk of dementia.

From the results of the multi-state model, I found that home care availability influences mortality and institutionalisation from patients own home, and they were spending less time inside hospital and therefore, their rate of in-hospital mortality was low. Increasing age, frailty, and hospital admission due to injury were associated with higher rate of

institutionalisation and death. Similarly, hospital stay ≥ 12 days was associated with hospital discharge to long-term care institutions and patients who were getting re-admitted within 30 days had a higher discharge rate and because of this had higher rate of rehospitalisation and subsequent institutionalisation and increased risk of death.

Conclusion and relevance

The gaps identified by this review will help researchers design better observational studies that will inform future trials more comprehensively. It also provides an alternative way to study the intricate dynamics of hospitalisation, institutionalisation and mortality using multi-state model and provide a foundation for further research with appropriate data on formal and informal home care. The work serves as a cornerstone for further research.

Impact Statement

The work in this thesis consists of a review of the literature on dementia outcome studies, using electronic health records data to understand dementia risk factors in diabetic patients, and using a multistate model to understand the intricate dynamics between hospitalization, institutionalization, and mortality in dementia patients.

I looked at the published literature in a systematic manner and highlighted the limitations and gaps in existing research, emphasizing the need for more comprehensive and multidimensional studies. By addressing these limitations, future research can provide more accurate information about how dementia progresses over time, leading to better-informed decisions and improved care for patients and their families. This highlights the need for more comprehensive and long-term studies that incorporate biomarkers, function, and quality of life. This review will guide future research. Researchers can use it as a roadmap for designing more robust observational studies and clinical trials, ensuring that future research provides a deeper understanding of dementia progression. This review can also help policymakers consider a multi-dimensional approach when developing policies and allocating healthcare resources. This can lead to more patient-centred care and improved resource allocation, ultimately enhancing the quality of life of patients with dementia.

The chapter on dementia risk factors in diabetic patients meticulously analysed a vast dataset over a 10-year period and unearthed crucial insights with far-reaching implications. The study directly benefits the public and patients with dementia and their families by providing a comprehensive understanding of the risk factors associated with dementia in patients with diabetes. Based on this knowledge, patients and their families can make

informed decisions regarding their healthcare, potentially delaying or mitigating the onset of dementia. The insights garnered here provide a solid foundation for researchers and clinicians to delve deeper into the intricate relationship between diabetes and dementia.

The use of a multi-state model provides a nuanced understanding of the intricate dynamics between hospitalization, institutionalization, and mortality in dementia patients and provides a deeper understanding of the factors influencing transitions and promotes interventions such as home care that can significantly enhance the quality of life of dementia patients. This study provides a rich foundation for further research endeavours in dementia care, and researchers can build upon these findings to explore nuanced facets of dementia patients' outcomes. These findings have the potential to shape health policies to directly benefit patients with dementia. Policymakers can optimize the allocation of limited healthcare resources by directing resources towards the most effective intervention and support system. The practical application of this study can lead to the development of targeted strategies aimed at reducing hospital readmissions and institutionalization.

In summary, the work in this thesis has a far-reaching impact that extends from the empowerment of dementia patients and their families to the optimization of healthcare policies and resource allocation. This serves as a cornerstone for academic research and offers a robust platform for further exploration. Through dissemination in scholarly journals, these findings have the potential to drive tangible improvements in dementia care, ultimately enhancing the lives of those affected.

Acknowledgements

I would like to express my sincere gratitude to the following individuals who have supported me throughout my PhD journey, despite the unique challenges I faced.
Dr Dionisio Acosta: Despite the limited supervision, your guidance and insights played a crucial role in shaping the direction of my research. Your input was invaluable, and I am grateful for your assistance.
I would like to acknowledge the contributions of Dr. Arturo Gonzalez who stepped in as my supervisor during a critical phase of my research. Your support and encouragement were instrumental in keeping me on track.
I am also thankful for the mentorship and advice provided by Dr. Paul Taylor.
Finally, I am proud of myself that during a very difficult phase of my life, I motivated myself and took the initiative to find relevant resources online and social media to complete my thesis.
Although my interactions were limited, each individual mentioned above played a vital role in my academic journey. Their contributions are greatly appreciated.

Table of Contents

Table of Tables

Table of Figures

Chapter 1: Introduction

This chapter is structured into three distinct sections. Section A delineates the aims and objectives of the study, establishing dementia as a paramount health concern. Section B provides a comprehensive overview of the thesis, succinctly summarising the content of each chapter. Finally, Section C delves into the intricacies of dementia as a syndrome, encompassing its manifestations, symptoms, subtypes, various diagnostic criteria, biomarkers, and the obstacles encountered in the development of novel pharmaceutical interventions for dementia.

1.1 Section A

1.1.1 Background

Dementia is a global health priority and as healthcare has become more efficient and patient-centric in developed countries, mortality rates have dropped and the number of people with dementia has increased. There are approximately 55 million people worldwide living with dementia. Taking care of dementia patients has a huge impact on resources. There is no primary prevention treatment available for dementia and pharmaceutical companies conducting trials to find a disease-modifying therapy have suffered huge setbacks as there is a very low success rate in these trials.

Understanding dementia outcomes and the factors affecting the rates of these outcomes is important for designing future studies and trials and will inform trials for future disease-modifying therapies. To understand these outcomes and factors affecting them, we need to first quantify longitudinal studies on dementia outcomes. We can then use statistical

modelling to estimate the hazards associated with different outcomes in dementia patients with different profiles.

Survival analysis is widely used in epidemiology and medical statistics and its most frequent use is in the form of the standard Cox proportional hazard model but there are other flexible parametric models such as Royston-Parmar model. This standard survival analysis involves only one event of interest and measures the time from some time origin until the occurrence of the event of interest. Typically, the outcome of interest is all-cause mortality or a composite outcome of death and non-fatal events. However, the event of interest could be death due to a specific cause, e.g., cardiovascular causes or there could be multiple events per subject such as repeated hospitalisation or infections episodes. These scenarios are called "competing risks" (CR) and require a different approach from traditional survival analysis. The CR model extends the Cox model by taking into consideration two or more events that subject is at risk of. The risks of these mutually exclusive events competing with each other to be the first to occur and hence the endpoint. So, one of the events precludes the other event. For example, if the outcome of interest is death due to cardiovascular cause, then death due to other causes serves as a competing event. Competing risks can also be present if the event of interest is non-fatal. For example, in case of discharge into a nursing home, discharge to patient's home is a competing event. Another example can be hospital discharge as a competing event when one is interested in in-hospital infection.

Standard competing risk time-to-event models are the simplest form of multi-state model. Such a model can be viewed as a multi-state model with one initial state and several mutually exclusive absorbing states. Multi-state models extend the competing risks model and also consider subsequent events after the initial one and therefore, can accommodate non-fatal events, which can be an intermediate state between the initial and final absorbing

state, all of which are of interest. Multi-state models (MSM) are stratified Cox models with covariates having a different effect for each transition. MSM is therefore capable of handling multiple transitions, covariate-specific effects and also overall transition probabilities using information about every transition, instead of just one transition within a process. MSM allow us to estimate absolute risk as the probability of being in a state over time and also to estimate dynamic predictions such as prediction of the probability of being in a state over time for a specific patient profile. This could have useful clinical applications.

1.1.2 Aims and Objectives

The aim is to describe the healthcare experience of people with dementia from the time they are diagnosed until their death using primary care and hospital episode statistics (HES) data.

- To quantify current knowledge of dementia progression through a systematic literature review so that future studies and trials are optimally designed and to identify gaps in current knowledge requiring further research.
- To characterize the identified dementia patients in electronic health records (EHRs) and their comorbidities before dementia diagnosis.
- To identify the factors associated with the risk of dementia diagnosis in the most occurring comorbid condition in dementia patients' pre-diagnosis.
- To use a multi-state model for hospitalized dementia patients for the joint analysis of outcomes and multiple hospitalizations, discharge home, discharge to a long-term care (LTC) institution and, mortality and estimate the effect of different covariates on the hazard rate of the different transitions. This is important because

discharge to a long-term care institution directly from hospital is not recommended and therefore, knowing the factors affecting this transition is important.

- To estimate the probabilities in different states of the model at some time point starting from the dementia diagnosis.
- And finally, to estimate predictions of probabilities in specific states at certain time points for specific profiles of patients, also called cumulative probabilities curves or predicted survival curves.

1.2 Section B

1.2.1 Structure of the thesis

Chapter 2 describes a systematic review of dementia outcomes in the literature. I characterize longitudinal cohort studies on outcomes post-dementia diagnosis and identify the outcomes they measure in these studies and identify opportunities to help inform future study and trial design.

Chapter 3 describes my data sources and explains the structure of the CPRD and HES data and how these data are selected for research. It then explains the selection of my cohort for the initial analysis from the CPRD data and estimates summary statistics about dementia patients in these data sources. I then identified comorbidities for dementia patients in HES data to compare with the comorbidities reported in the literature review of dementia outcomes studies.

Chapter 4 describes the hazard rates of dementia diagnosis in patients who did not have dementia at first but were diabetic to start with. From the results of the literature review and the HES data, diabetes was the top occurring comorbidity in

dementia patients, and I wanted to see the effect of diabetes on getting diagnosed with dementia. I performed a time-to-event analysis using a Cox proportional hazard model to see what the main factors in diabetic patients for dementia diagnosis were.

Chapter 5 introduces multi-state modelling. In chapter 4 I looked at the factors affecting dementia diagnosis in diabetic patients using the traditional Cox model. In this chapter, I describe multi-state modelling approach which provides the scope for the dementia outcomes post-diagnosis in the following chapter.

Chapter 6 presents a Markov Multi-state modelling of dementia, recurrent hospitalisation, institutionalisation, and death. The results of the effects of covariates on all relevant transitions between the different states are shown and interpreted. I also illustrate the estimation of cumulative probabilities curves for a specific profile of patients which shows the probability of a state considering the time elapsed since dementia diagnosis.

1.3 Section C

1.3.1 Dementia as a syndrome

Dementia is a devastating neurodegenerative disease with dire implications for individual patients, their relatives, carers, the healthcare systems, and society in general. It is a long-term disease characterized by gradual problems over time with cognition (memory loss, executive function), daily functioning and behaviour [1]. Dementia is a syndrome, and it is not part of normal ageing but is a collection of diseases that affect the brain due to loss of nerve cells. It is known that depression, delirium, thyroid problems, certain vitamin

deficiencies, side-effects of some drugs , smoking and excessive consumption of alcohol cause dementia-like symptoms [2].

1.3.2 Epidemiology of Dementia

Dementia is a syndrome characterized by a decline in cognitive function that interferes with daily life. It primarily affects older adults, although it can occur in younger individuals. The prevalence of dementia increases with age, with estimates suggesting that around 5-7% of individuals over the age of 60 are affected worldwide [3] . Alzheimer's disease is the most common cause of dementia, accounting for approximately 60-70% of cases (Alzheimer's Association). Currently more than 55 million people have dementia worldwide, over 60% of whom live in the low-and middle-income countries. Every year there are nearly 10 million new cases added.

1.3.3 Inequalities in Dementia

There are significant inequalities in the prevalence, diagnosis, and management of dementia across different populations. Socioeconomic status, education level, and access to healthcare services play crucial roles in determining the risk of developing dementia and the quality of care received [3]. Minority ethnic groups and marginalized communities often experience disparities in dementia care, facing barriers such as stigma, cultural beliefs, and language barriers [4].

1.3.4 Changes Over Time

The prevalence of dementia is expected to increase significantly in the coming decades due to population aging and changes in lifestyle factors such as diet, physical activity, and smoking habits . Advances in medical technology and diagnostic criteria have also led to improved detection and reporting of dementia cases over time, contributing to apparent increases in prevalence rate [3] .

1.3.5 Differences Across Countries and Regions

There are notable differences in the prevalence and management of dementia across countries and regions. Developed countries with aging populations tend to have higher prevalence rates of dementia compared to developing nations [3]. Variations in healthcare infrastructure, access to resources, and cultural attitudes towards aging and cognitive decline influence the burden of dementia in different regions (Alzheimer's Disease Internat.

1.3.6 Cost of Dementia

Dementia imposes a significant economic burden on individuals, families, and healthcare systems. The direct costs of medical care, including diagnosis, treatment, and long-term care, are substantial. Additionally, informal care provided by family members and caregivers represents a significant but often unrecognized cost of dementia . In 2019, dementia cost economies globally 1.3 trillion US dollars, approximately 50% of these costs are attribute to care provided by informal carers (family member and close friends), who provide on average 5 hour of care and supervision per day [5].

1.3.7 Management Strategies

Management of dementia involves a multidisciplinary approach aimed at optimizing cognitive function, managing behavioural symptoms, and providing support for patients and caregivers. Pharmacological interventions, including cholinesterase inhibitors and memantine, are commonly used to manage cognitive symptoms [6]. Non-pharmacological interventions such as cognitive stimulation therapy, physical exercise, and psychosocial support programs are also essential components of dementia management [7].

People with dementia can take steps to maintain their quality of life and promote their ell-being by (1) Being physically active and (2) Maintain social interactions that stimulate

the brain and maintain daily function. Providing care and support in their usual place of living to dementia patients has great potential to increase their quality of life, and decrease mortality.

1.3.8 Types of dementia and causes

Alzheimer's dementia (AD), which is the most common subtype accounting for more than 60% of all the cases, is caused by Alzheimer's disease which is characterized by cortical amyloid plaques and neurofibrillary tangles while vascular dementia (VaD), which is the second most common subtype, accounting for about 20% of cases, and diagnosed when the oxygenated blood to the brain is disrupted after a stroke or other blood vessel complications [4]. However, there is not a very clear distinction between AD and VaD, probably because vascular complication is just a co-factor in people with Alzheimer's disease, enhancing clinically significant symptoms and therefore, about 22% of elderly have pathology of mixed vascular-Alzheimer's dementia (MVAD) [10]. Other less common subtypes are dementia with Lewy bodies (DLB), Parkinson's dementia (PD) and fronto-temporal dementia (FTD). Some of the main types and their characteristics are described in Table 1.

Table 1: Different types of dementias and their characteristics.

Adapted from Alzheimer's Association- 2018 Alzheimer's disease facts and figures, Deramecourt et al, 2012 and Kalaria, 2016[2,11,12]

Cause	Prevalence	Onset	Early clinical symptoms	Later symptoms	Pathology
Alzheimer's disease	60 – 80%	Slow onset and progression	• Difficulty remembering things. • Apathy • depression	• Impaired communication • Disorientation, confusion • Behaviour change • Poor judgement • Difficulty in activities of daily living	• Beta amyloid (plaques) outside neurons in brain • Tau tangles inside neurons
Vascular dementia (VaD)	~ 20%	Can be a gradual or an abrupt progression	• Impaired judgement • Impaired ability to organize things	• Problems in memory and cognition • Difficulty with motor functions such as gait, balance	• Blood vessel blockage • Ischemic or hemorrhagic infarcts • White matter changes
Dementia with Lewy bodies (DLB)	15%	Insidious	• Sleep disturbance • Visual hallucination • Gait imbalance	• Cognitive impairment	• Lewy bodies (abnormal aggregation of alpha-synuclein proteins in neurons. • Coexisting Alzheimer's pathology

Frontotemporal dementia (FTD)	<5%	Early, insidious onset	• Marked changes in personality & behaviour. • Difficulty in comprehending language.	• Similar symptoms to Alzheimer's • 60% FTD occur at age 45 to 60.	• Nerve cells in the frontal and temporal lobes shrunken. • Upper layer of cortex become soft and spongy.
Mixed dementia	22%	Insidious	• Symptoms depend on the types of dementias. • Usually, Alzheimer's and VaD followed by Alzheimer with FTD		• Characterized by hallmark abnormalities of more than one cause of dementia- Mostly Alzheimer with Vascular, followed by Alzheimer's with FTD

1.3.9 Diagnosing dementia

Dementia is one of the main reasons for dependence and disability [13,14] and therefore, a timely diagnosis can help carers/family members and healthcare providers improve the management of the disease and also resource allocation.

1.3.9.1 Dementia in primary care

General practitioners (GPs) are in an ideal position to spot any symptoms of early dementia and can refer the patients for further investigation [15].They can also exclude other treatable illness or reversible causes of dementia such as Vitamin B12 deficiency, thyroid problems or any acute depression problems and refer to specialists or memory clinics in case of neurological or behaviour or more complex cases.

Assessment in primary care for diagnosing dementia includes taking patient's history to understand any changes in cognition or behaviour and the dependency of the patient on carers in daily life. Further assessments involve a blood test, brief cognitive assessment such as GP assessment of cognition (GPCOG), which is a five minute questionnaire where a score of >8 means an impairment in cognition while <5 means that the cognition is still intact [16].Other cognitive test include Mini-Cog [17] , which consists of two components: a three item recall and the clock drawing test. Cognition is considered impaired if the patient cannot recall any of the three items or they can only recall one or two things and draw an abnormal clock [13]. It is recommended to carry out these brief cognitive assessments before referral to secondary care [18].

GP can refer the patient to specialist dementia service including appropriate neurological examination and cognitive and functional testing.

1.3.9.2 Diagnosis in Secondary care

From 2012, with the recommendations from the department of health every person aged 75 and over goes through screening for dementia in hospital[19]. Patients or their carers/family members are asked about any changes in memory or functioning of the patients. This is followed by cognitive tests, physical examination, body fluid tests to identify conditions causing the symptoms. They also arrange brain scans and specialist assessment by psychiatrists when needed [19,20].

Diagnoses of dementia in the hospital episode statistics data (HES) are clinically identified, obtained from correspondence with primary care, or from existing hospital records as the system pre-populate diagnosis field from previous record of the patient chronic conditions [19]. The accuracy of hospital diagnosis of dementia has been reported to have a sensitivity of 78% and specificity of 92% [19]. Other studies have reported a specificity of 98% and 99% [21,22].

1.3.9.3 Diagnosing dementia subtypes

Dementia subtypes can be diagnosed after the initial cognitive and neurological examination, using the validated and international standardised criteria. Biomarkers are part of the new standardised guidelines for dementia and its subtypes diagnosis and for understanding the severity of the disease.

International consensus criteria for dementia with Lewy bodies

International criteria for frontotemporal dementia

NINDS-AIREN criteria for vascular dementia

International Classification of Disease (ICD)

ICD, which contains guidelines for recording and coding of health conditions, is published by the World Health Organisation (WHO). ICD translate diagnoses of diseases and health problems from words into an alphanumeric code. It is an international standard diagnostic classification system for all general epidemiological and health-management purposes. It is used in the UK secondary care and also in the mortality registry data.

National Institute of Ageing- Alzheimer's Association (NIA-AA)

The NIA-AA new guidelines aim to improve he diagnosis by incorporating biomarkers to strengthen the diagnostic accuracy. Details can be found on in the link https://www.alz.org/research/for_researchers/diagnostic-criteria-guidelines

1.3.9.4 Alzheimer's disease biomarkers

A biomarker can be defined as a characteristic or substance that can be measured and evaluated and acts as a predicator of a biological process, disease or pharmacologic response to an intervention [23].With the prospect of disease-modifying treatment, biomarkers have been the focus of research for early diagnosis and dementia subtype identification [24,25].

The recent "AT(N)" classification recognizes three groups of biomarkers [26]. In this classification Ab plaques are denoted by "A", neurofibrillary tau is denoted by "T" while the "N" denotes neurodegeneration. The neurodegeneration is not Alzheimer's specific and that's why its enclosed in the bracket.

1.3.9.5 Cerebrospinal fluid (CSF) biomarkers

1.3.9.5.1 CSF Aβ

It is now widely accepted that CSF- Ab42 is a major component of senile plaques and a valid biomarker of cerebral Ab pathology in Alzheimer's disease and dementia [27,28]. Ab42 is formed when the transmembrane amyloid precursor protein (APP) is cleaved by b- and g- secretase. The CSF concentration of Ab42 is inversely related to the degree of amyloid plaques in the brain [29]. The CSF sample can be collected by lumbar puncture which is a very invasive procedure and the concentration of Ab42 can be measured by a technique called "enzyme-linked immunosorbent assay" or ELISA, and also by mass-spectrometry [30]. Studies have demonstrated that Ab42 concentration reduces by ~50% in Alzheimer's disease patients when compared to individuals of the same age with no Alzheimer's disease [29,31].

1.3.9.5.2 Amyloid-PET

Positron emission tomography (PET) and radiotracer ligands with high affinity for Ab are used to provide evidence of amyloid plaque load in the brain. In amyloid-PET, radiotracer ligands such as C^{11} -labelled modified derivatives of amyloid-binding histological dye thioflavin-T also called as Pittsburgh compound-B is used with PET (C^{11}] PiB-PET) to measure amyloid load [32]. Other ligand compounds such as F^{18} -flutemtamol (vizamyl) and F^{18}-florbetaben (Neuraceq) have a longer half-life than C^{11} and are used in PET to detect amyloid burden [32].

1.3.9.6 Tau biomarkers

1.3.9.6.1 CSF p-tau

Neurofibrillary tau (denoted "T" in the AT(N) classification) is a protein that stabilizes the microtubules in the neuronal axons but when they are hyperphosphorylated (p-tau), they become abnormal and form tangles inside neurons [32,33]. The CSF of Alzheimer's disease individuals are known to have an elevated level of p-tau which can be measure by ELISA and is considered one of the specific biomarker for Alzheimer's [29,34,35].

1.3.9.6.2 Tau-PET

Tau-PET is a new modality[36] and the current ligand compounds have some limitations in their binding and research for more compounds which can bind more reliably to tau is underway [37]. One such tracer is the F^{18} - AV1451 and has shown its affinity to tau- tangles and can differentiate Alzheimer's brain from healthy controls [38,39]. The elevated tau-PET binding in the brain is strongly associated with positive amyloid-PET and clinical symptoms of dementia [40,41].

1.3.9.7 Neurodegeneration/Neuronal injury (N) biomarkers

In the new AT(N) system for Alzheimer's disease biomarkers classification, neurodegeneration biomarkers are denoted by N and because neurodegeneration is not Alzheimer's-specific and can be caused by any other etiology such as cerebrovascular injury [42], it is therefore, placed in parentheses. Biomarkers for neurodegeneration include magnetic resonance imaging (MRI), fluorodeoxyglucose (FDG-PET) imaging and CSF concentration of total tau (t-tau) [36].

1.3.9.8 CSF total tau (CSF t-tau)

Total tau (t-tau) level in the CSF of Alzheimer's patients has been observed and it is also considered a biomarker for neurodegeneration [29]. This increase in t-tau level is not specific to Alzheimer's disease and the increase is seen in traumatic brain injury and stroke which correlates to neuronal injury [43–45]. Similarly, in Creutzfeldt-Jacob disease, there is a big increase in t-tau level but no change in p-tau [46,47] . However, in Alzheimer's disease there is a constant increase in p-tau [29] reflecting the abnormal Alzheimer's pathology associated with tau tangles and CSF t-tau reflects neurodegeneration and therefore, the levels of p-tau and t-tau along with an abnormal MRI and amyloid pathology can be very strong diagnostic criteria for Alzheimer's dementia [36,48,49].

1.3.9.9 FDG-PET

Fluorodeoxyglucose PET (FGD-PET) is an example of functional brain imaging and can reveal how well the cells in various regions of the rain are performing by showing how actively they utilize oxygen and sugar and therefore, detect brain region-specific impairment [50]. FDG-PET can add diagnostic accuracy because if the result is a reduction in glucose metabolism then it indicate neuronal injury [51].

1.3.9.10 Structural imaging

Structural imaging provides information about the shape, position or volume of brain tissues [52]. Structural imaging such as computed tomography (CT) or magnetic resonance imaging (MRI) are used in dementia diagnosis. MRI provides much better resolution than CT, and is more sensitive to subtle vascular changes and therefore, used in differential diagnosis ,e.g., Fronto-temporal dementia (FTD), multiple sclerosis [51]. Structural MRI can therefore provide a measure of brain atrophy, tau tangles load [53,54].

1.4 Current drug treatments for dementia

Since there is no disease-modifying treatment (DMT), the currently available drugs to dementia patients only help to alleviate the symptoms with limited benefits. The two classes of drugs available for cognitive symptoms are:

(i) Acetylcholinesterase (AChE) inhibitors

(ii) Anti-glutamatergic

1.4.1 Acetylcholinesterase (AChE) inhibitors

Acetylcholinesterase inhibitors block the breakdown of acetylcholine by the enzyme cholinesterase. Acetylcholine is a neurotransmitter and thus helps in neuronal communication. AChE inhibitors currently available on the NHS are donepezil, rivastigmine and galantamine. Memantine is another anti-dementia drug and belongs to the anti-glutamatergic group which helps in reducing the neurotoxic effects of excessive glutamate released [55].

1.4.2 Drugs for behavioural symptoms

Psychotropic drugs are used to treat behavioural and psychological symptoms (BPSD) of dementia. In dementia and other neurogenerative diseases there are structural and chemical alterations which result in symptoms such as BPSD and antipsychotics can help to control these symptoms such as aggression, delusion, hallucination, depression, and apathy [55].

1.5 Development of new drugs for dementia

As of January 30, 2018, there were 112 agents in the Alzheimer's disease treatment pipeline[56]. The percentage of compounds classed as disease modifying treatment (DMTs) were 63% (phase I through III) and the rest were cognitive enhancers or targeting behavioural problems [56]. Dementia drug development has been very difficult with a 99.6% failure rate since 2002 [57] and there has been no new medication for dementia approved by FDA since 2003. There were 26 agents in phase III in the 2018 drug pipeline and 17 of them were DMT's [56]. However, so far there has been a 100% failure of disease modifiers' phase III trials which are also the most costly part of the dementia drug development costing $413 million [58].

1.6 Potential treatments

1.6.1 Amyloid beta (Aβ) plaques

The build-up of Amyloid beta (Ab) plaques in the brain is a key feature of Alzheimer's disease [59]. A number of trials are targeting the Ab and in the 2018 Alzheimer's drug pipeline, there were 26 agents in phase III, and 17 of them were DMTs. Out of these 17 DMTs, 14 were for amyloid target, 1 involved a tau related target, 1 for neuroprotection and 1 other had metabolic mechanism of action (MOA) [56]. However, most of the recent phase II/III trials with promising drugs have failed such as verubecestat from Merck & co and solanezumab by Eli Lilly [60]. More recently scientists at the University of Southern California (USC) in Los Angeles have explored a new compound called the "3K3A-APA", which is genetically modified version of the activated C protein found in the blood to protect blood cells and vessels from damage due to inflammation [61]. In tests, a lowering of Amyloid- b was observed in genetically altered mice injected with 3K3A-APA and it

has been shown to prevent brain cells from making the enzyme BACE1 which is essential in Amyloid- b production [62].

Similarly, a new therapeutic antibody candidate for Alzheimer's treatment is PMN310, which is in late preclinical development and has the ability of targeting Amyloid- b and is anticipated to enter phase I trials in mid 2019 [63].

1.6.2 The immune system

Another treatment area is using anti-inflammatory drugs with potential to treat the inflammation in the brain triggered by the immune system in dementia [64,65]. Two drugs already in the market for different conditions (Pioglitazone for diabetes and Etanercept for arthritis) were promised in phase II trials but unfortunately the trials were ended in phase III stages [66,67].

1.6.3 Tau tangles

Tau is a microtubule-associated protein which in its hyperphosphorylated form make tau tangles, one of the hallmarks of Alzheimer's disease [68]. These intracellular tau aggregates are common in a number of neurodegenerative diseases, including Alzheimer's and are associated with synaptic loss and neuronal death [69]. Some of the DMTs targeting tau pathologies are focusing on protein kinases as the tau hyperphosphorylation is due to the action of protein kinase and therefor, several protein kinases are included on the therapeutic pipeline [70,71]. Apart from phosphorylation, tau is also modified post-translationally by lysine acetylation and which leads to abnormal tau functionality promoting tau aggregation and therefore, tau acetylation inhibitors are another potential therapeutic strategy [72]. Nonsteroidal anti-inflammatory drugs (NSAID) which are used for inflammatory conditions such as arthritis have shown effectiveness against tau-acetylation and improved neurodegeneration in mouse models [73].

1.7 Challenges in developing new drugs for dementia.

There are many challenges in developing anti-dementia drugs.

(i) Because of the blood-brain barrier which prevent some drugs from reaching it

(ii) Because of the nature of dementia as a disease caused by different disease and involve multiple pathologies such as Ab and tau tangles and also the lack of a surrogate biomarker [74,75].

(iii) Currently, there are Alzheimer's disease biomarkers which can improve diagnostic accuracy but no accepted surrogate biomarkers [74,76].

DMTs will act by affecting the underlying dementia pathologies such as Ab or tau [77,78]. The concentration and burden of Ab and tau can be measured from CSF or using brain imaging [79], however, more research is needed to find a surrogate biomarker which can be substituted for a clinical end point [74]. As DMTs will target underlying dementia pathologies, we will need participants who definitely have dementia pathologies to assess the effect of the drug [59].

With the above challenges in drug development and a failure rate of close to 100% [57] and the predicted tripling in dementia prevalence by 2050 [80] , improving the design of clinical trials and dementia diagnostics are urgently needed.

Longitudinal cohort studies have been an important source of information for clinical trials [80]. DMTs will have to be validated in well-conducted clinical trials, however, as we know that conducting trials in a slowly progressing disease is costly and challenging in terms of the number of participants [80]. It is, therefore, very important that we fully understand the dementia disease progression in a representative cohort and for that the role of longitudinal cohort studies is crucial. Therefore, I conducted a systematic review of longitudinal cohort studies on dementia progression in order to find what these studies are focusing on and what can we learn from them to help us design better .

Chapter 2: A systematic literature review of dementia outcome studies

2.1 Introduction

Dementia is a global health priority. In 2018, there were approximately 53 million people worldwide living with dementia, expected to increase to 75 million by 2030, and the total estimated societal cost was US$ 1 trillion in 2018 [81]. In the US alone an estimated 5.7million people were living with dementia in 2018 [2]. There is no primary prevention treatment available, and the current interventions are secondary or tertiary in nature, largely focused on alleviating the symptoms. Acetylcholinesterase (AChE) inhibitors are the only group of medications to manage and delay the progression of dementia symptoms while no effective specific medication exists for vascular dementia beyond existing cardiovascular therapies [82]. Given the increase in prevalence and cost to the economy, novel disease-modifying treatments (DMT) which will prevent or delay the onset or slow down the progression are urgently needed.

According to a new report by the Pharmaceutical Research and Manufacturers of America (PhRMA), 146 investigational medicines in clinical development were halted between 1998 and 2017 [83] . The report also shows that the large portion (39%) of these drugs were in phase 2 or 2/3 and a substantial (18%) were in phase 3 or regulatory review. Only four medicines have been approved by the US food and drug administration (FDA) since 1998 which comes to about 1:37 ratio of success to failure and represents about 2.7% success rate to date. Longitudinal cohort studies are ideal for understanding the multifactorial and slowly progressive nature of a disease such as dementia and therefore, represent an important resource of information for designing clinical trials [84]. Therefore, I conducted a systematic literature review of longitudinal cohort studies to understand dementia progression from multiple dimensions covering clinical, biomarker and health system utilisation perspective. The objective was to identify the different outcome measures used,

highlight limitations, any unexplored areas and identify opportunities to help inform new observational studies and clinical trials.

2.2 Methods

2.2.1 Data source and search

A literature search was conducted in OVID MEDLINE from April 2008 till April 2019. The search strategy included the relevant synonyms for (1) Dementia/Alzheimer's (2) Disease progression (3) Health system utilisation (Supplements).

2.2.2 Inclusion/exclusion criteria

I included longitudinal cohort studies (prospective, retrospective) following individuals diagnosed with dementia and aged 65 or over. Studies focusing on rare dementias (such as frontotemporal or Parkinson's dementia), conversion of mild cognitive impairment to dementia or studying response to drug intervention were excluded. Studies were also excluded if focusing solely on the monetary or economic aspect or focused on animals or not in the English language. The search and selection process are described in figure 1. The selection was a consensus process between three reviewers KK, JP, and CD.

2.2.3 Data Extraction

Data from the methods and results sections of the selected studies were extracted by a single reviewer into an excel sheet (Table e1) and were checked twice for accuracy. Data were extracted relating to (1) study design (2) demographics (location, sample size, age, gender ratio) (3) follow-up duration and repeated measurements (4) diagnostic criteria (5) dementia subtypes (6) outcome measures of different domains such as clinical, biomarker, and types of health resources utilised.

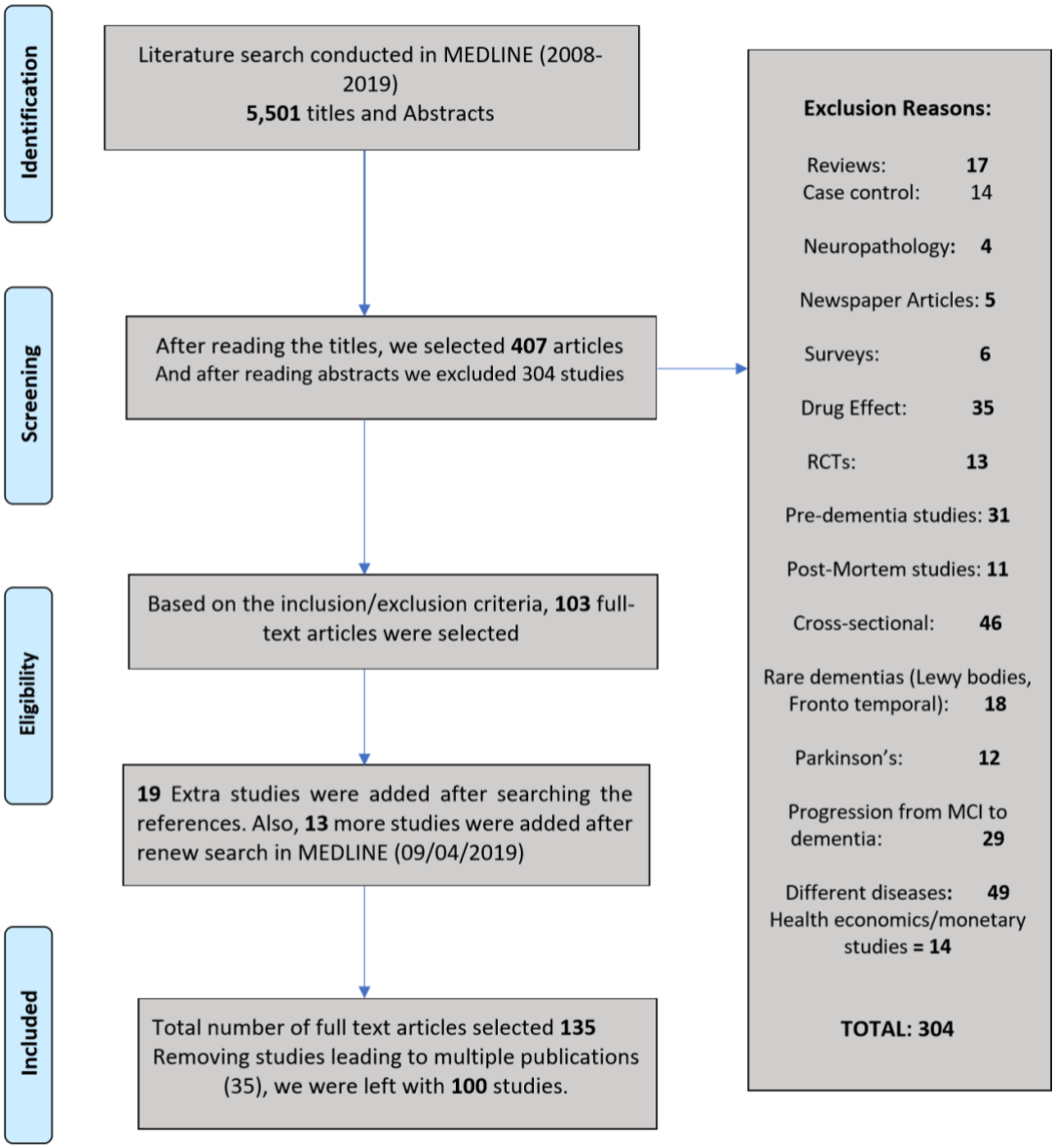


Figure 1: PRISMA flow diagram of the search and selection process

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal. pmed10000

2.3 Results

I screened 5,501 titles and abstracts out of which 407 articles were selected for further review. Of these, 116 articles fulfilled the inclusion criteria, and 19 other studies were added following references from the selected studies. Where a study leads to multiple publications, only the one with the largest sample was included. This resulted in the exclusion of 35 papers, and therefore 100 studies were selected in this review comprising >2m individuals (Fig 1).

2.3.1 Geography

I observed a large disparity between the global prevalence of dementia compared with where the selected studies in the review were conducted. There was a bias towards North America and Europe. For details see fig e1 & e2. Most of the selected studies originated from Europe (46) in which 9 were from the UK, Americas (31), and Scandinavia (12) while 11 studies were from the rest of the world (Fig e3).

2.3.2 Sample size (N) and follow-up

Of the 100 studies, 82% had N <= 5000, 13% had N>5000 and only 5% with >50,000 sample size. The follow-up of 66% of studies was <=3 years while only 4% had follow-up >10 years (Fig e4).

2.3.3 Types of dementia investigated.

I looked at the reported prevalence of dementia subtypes in the selected studies and found the studies dominated by AD. In the selected studies, 85% of the total sample size (N) was of AD and only 9% was VaD while 6% were unclassified or other subtypes. When the dementia subtypes reported in the studies in my review were compared with a study of dementia subtypes recorded in the English Electronic Health Records (EHRs), using the CALIBER (CArdiovascular disease research using Linked Bespoke studies and Electronic Health Records) platform, this compared to 27% as AD, 20% as VaD and 53% unclassified [85] (Fig e5).

These figures for dementia prevalence in primary care are much lower than the prevalence reported in the general population which states that AD is the more common subtype corresponding to more than 60% and VaD for ~ 20% of all the cases [8,9,86]. A study conducted in the UK using the expert Delphi methodology [86] estimated a similar prevalence. The difference in prevalence in the primary care data available in the electronic health records could be partly because there are low recordings of specific dementia subtypes in the primary care and most of the cases are recorded as unclassified. As a result, dementia subtypes such as VaD are underrepresented in the studies in the research to date and the studies are mostly dominated by AD patients.

2.3.4 Diagnostic methods and sample selection

I considered which diagnostic criteria or diagnostic codes were commonly used in the selected studies. I observed a good degree of consensus on this as more than 70% of studies using the revised National Institute of Neurological and Communicative Disorders and Stroke (NINCDS-ADRDA) diagnostic criteria either on its own or in combination with the Diagnostic and Statistical Manual (DSM-III/IV), National Institute of Ageing- Alzheimer's Association (NIA-AA), National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et I'Enseignement en Neurosciences (NINDS-AIREN), or biomarkers and brain imaging [87–90]. Figure 2 shows that the studies with larger sample sizes have weaker diagnostic or sample selection criteria as they were mainly from insurance databases and using ICD codes, mostly without any validation or additional codes. The diagnostic criteria were classed as weak when only diagnostic codes without validation / or without other associated codes were used and optimal/strong when one of the standardised diagnostic criteria along with biomarkers or brain imaging were used.

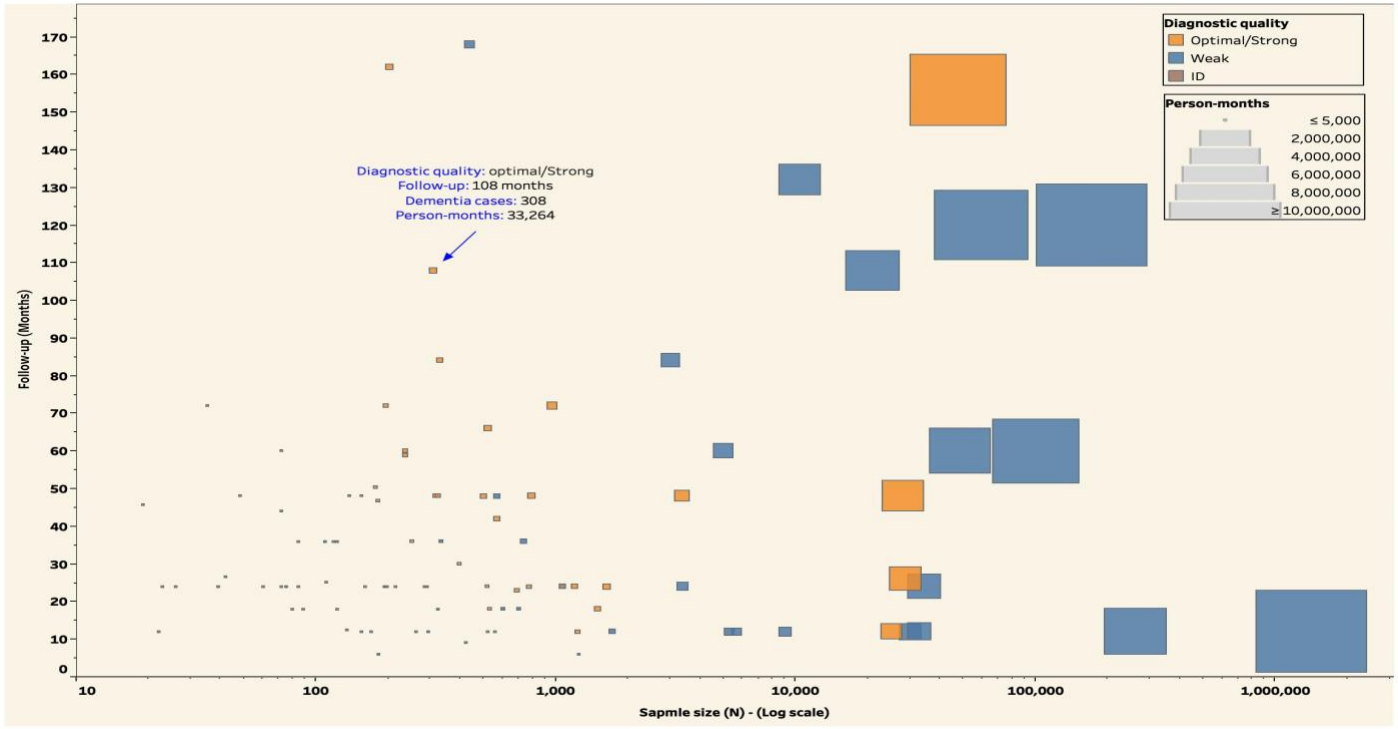


Figure 2: Association between sample size, follow-up duration and quality of diagnostic criteria.

Optimal /strong: (a): NINCDS-ADRDA, NINDS-AIREN, on their own or with NIA-AA or DSM-III/IV
(b): (a) in different combinations or with histopathological confirmation
(c) ICD codes with prescriptions or death certificate validation
Weak: (a): DSM-III/IV or ICD codes on their own with no validation
(**b)** MMSE, CDR or other neurological assessments on their own.
ID: Insufficient data
x-axis represents sample size (N) and y-axis the average follow-up (months) of the studies.

2.3.5 Brain imaging and other biomarkers

Biomarkers are characteristics that can be measured and evaluated objectively as an indicator of normal biological processes, pathogenic processes or pharmacologic response to a drug intervention [23] . There are different varieties of dementia biomarkers: Cerebrospinal fluid (CSF) amyloid β 1-42 and molecular brain imaging such as Pittsburgh compound B- PET [C^{11}] PiB-PET are sensitive to disease presence but not necessarily progression, similarly, there are biomarkers which are sensitive to disease progression but not necessarily its presence such as MRI, and fluorodeoxyglucose (FDG-PET) positron emission tomography which can detect neurodegeneration but are not specific for AD and rather a non-specific indicator of damage which may be because of other aetiologies such as cerebrovascular injury [25,36,42,91–95]. Genetic biomarkers include autosomal mutations in the Amyloid precursor protein (APP), presenilin -1 (PSEN-1) and PSEN-2 located on chromosome 21, 14 and 1 respectively and are linked to the familial Alzheimer's disease (FAD) [96]. While apolipoprotein (ApoE ε4) is a major genetic risk factor for sporadic or late-onset Alzheimer's dementia and is associated with early dementia and also a rapid decline and a higher [C^{11}] PiB-PET uptake [97–99]. PiB-PET is a specific positron emission tomography and used to identify Amyloid and has an affinity for Amyloid plaques and when used with PET can help in studying the progression of the Amyloid load [100]. Research has shown that the combination of low CSF Amyloid β-42 and high tau levels are associated with Alzheimer's dementia pathologies on post mortem with very high accuracy [101–104]. In the literature review, 20 studies were reporting Alzheimer's biomarkers, but the majority used it mainly for diagnostic purposes with measurements taken only at baseline and did not report changes that occur over time. Only three studies reported a longitudinal measurement of biomarkers. One reported the use of Pittsburgh compound B Positron Emission Tomography (11C PiB-PET) at

baseline and then at 24 months interval to investigate the association of amyloid load with APOE allele presence [105], while the second study [106] reported the results for cohorts with three or more assessments of PiB-PET during the follow-up to calculate the rate of amyloid deposition, cerebral atrophy, and cognitive decline. The third study investigated the longitudinal trajectories of CSF biomarkers for neurodegeneration (t-tau and Neurofilament light (NfL)) and tau pathology-associated biomarker p-tau and another novel marker called the YKL-40 which is an astrocytic marker [107]. They found that NfL and YKL-40 levels elevate longitudinally in AD and with age but not the CSF tau over a median lumbar puncture interval of 2.1 years. NfL levels in CSF reflect axonal damage and correlates to disease progression and brain atrophy [108]. YKL-40 is an inflammatory marker protein expressed in the astrocytes in the brain and its level also increases in early AD [109–111]. The results are summarised in figure e6. In summary, there was a clear lack of studies measuring biomarkers longitudinally.

2.3.6 Measures of progression

In total 52 studies (52%) reported repeat measurements of dementia-specific outcomes. Out of these 52 studies, 31 (~60%) measured them at 1 to 3 (median 3) separate time points after baseline measurements (Fig 3). More importantly, these measurements were taken within a time frame of just under 3 years (median 2 years). This is not long enough follow-up for a disease with a long prodromal development often lasting decades [112]. Dementia progression is multidimensional, and a single type of measurement is unlikely to capture adequately the different dimensions of the disease [113,114]. I looked at what other outcomes were measured along with cognition. For details see Fig 4. It was found that studies mostly measured cognition and dementia-specific outcomes were not measured longitudinally over a long enough period.

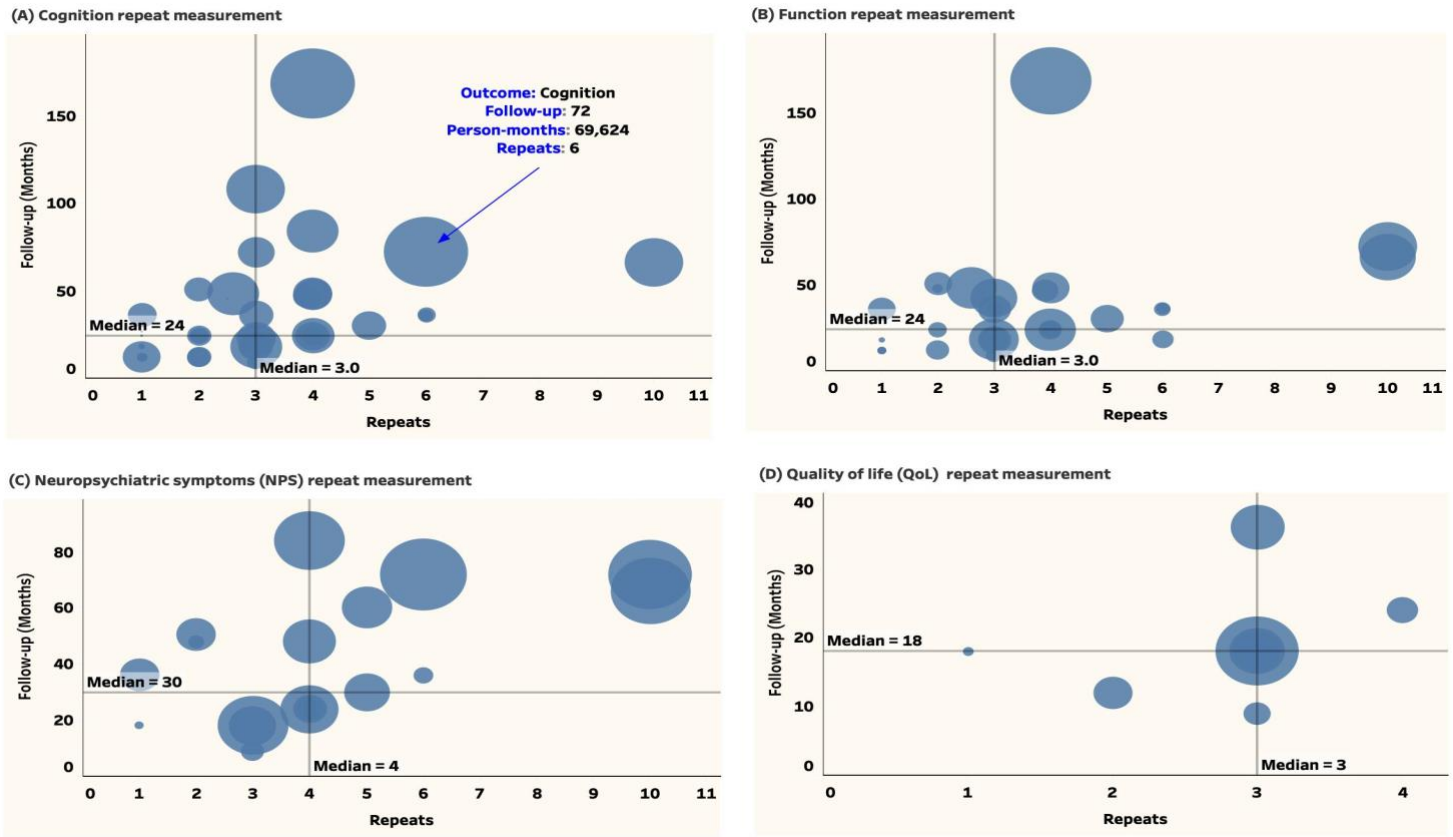


Figure 3: Scatter plots to illustrate the association between repeat measures of outcomes and length of follow-ups.

X-axis represents the number of repeated measurements (assessment occasions) after the baseline measures and the Y-axis is the follow-up duration in which these repeats were taken. For example, the annotated circle in (A) mean that the study measured cognition 6 times during a 72-month follow-up, and it corresponds to >69000-person months. Multiple Circles (darker colours) occupying the same position means multiple with the same repeat and follow-up. The size of the circles represents the person-months (N* follow-up(months)).

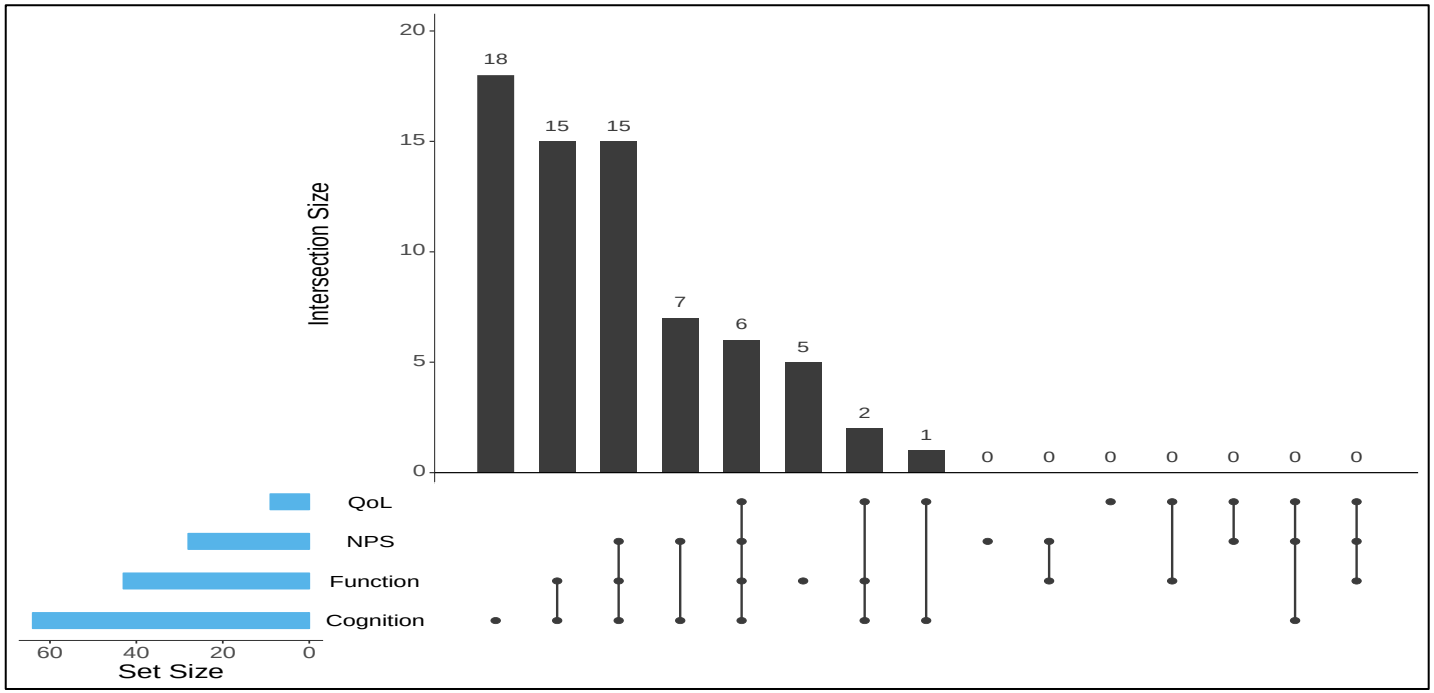


Figure 4: Diagrammatic representation of overlaps in the reporting of cognitive, functional, Neuropsychiatric (NPS) and Quality of life (QoL) outcome measures.

The above chart shows that cognition was the most measured outcome (total 69 studies measured cognition).18 studies measured cognition alone, while 15 measured cognitions + function, another 15 studies reported measuring cognition + function+ NPS while only 6 studies measured all the four outcomes simultaneously.

2.3.7 Electronic health record (EHRs) and administrative databases studies

In total 26 studies used health care data stored in electronic databases but only 3 of them were using linked EHRs while the rest of the studies used health insurance databases to identify dementia patients using the diagnostic codes (ICD, read codes). Out of the three studies using linked EHRs , one used "The Health Improvement Network" (THIN) from the UK, with >30k dementia cases, to estimate the rate of strokes in patient with and without dementia in the UK [115] while two used Clinical Practice Research Datalink (CPRD) from the UK : one reported the association of comorbidity burden in individuals diagnosed with dementia with resource utilisation and mortality [116]. They found that people with dementia and a higher number of comorbid conditions die early and have higher health resource utilisation. However, the study had a small sample size with N < 5000 and reported a limited number of comorbidities and lacked clinically validated dementia subtypes. The second study used the CALIBER platform which is a linked dataset of CPRD, hospital episode statistics (HES) and mortality data [85]. The authors aimed to determine the diagnostic validity of dementia captured in the EHR and to determine their lifetime risk of dementia. They found that the majority of people with a record of dementia had corroborating evidence of diagnosis and that the risk of dementia was higher in women than men and mortality was higher for people with than without dementia.

In summary, the studies using electronic health records were mainly using administrative databases and only a few studies used linked EHRs and therefore caution must be taken when interpreting the results from these databases.

2.3.8 Co-morbidities

Comorbidities are common in dementia patients and are one of the key factors in the higher utilisation of resources and risk of death [117,118]. Understanding their prevalence and association with the utilisation of different health systems and mortality rate is an important factor in planning resource allocation and for describing heterogeneity in the clinical trajectory of dementia based on comorbidity profiles. There were 35 studies (N = 2,008,535) reporting prevalence of comorbid conditions. The weighted percentages of the prevalence of the comorbid conditions reported in these studies are described in figure 5, which shows that diabetes and cardiovascular and cerebrovascular conditions were the top comorbid conditions while other symptoms and complications of dementia such as blindness, constipation and chronic pain were less reported in the included studies. The reason could be that these symptoms are recorded as free text and usually not coded for many patients and would require natural language processing (NLP) to extract this information [85]. Another, important aspect of the studies reporting resource utilisation was the use of administrative databases for their sample selection and had larger sample sizes and contributed ~84% of the total sample size (N) of all studies. As the primary aim of these databases was payment and operations, caution is needed when interpreting their results. I also found that although cognition and function were measured in most of the studies, the sum of sample size in these studies was very small (Fig e7) which means only a fraction of the total sample in the studies in this review underwent dementia-specific outcome measures.

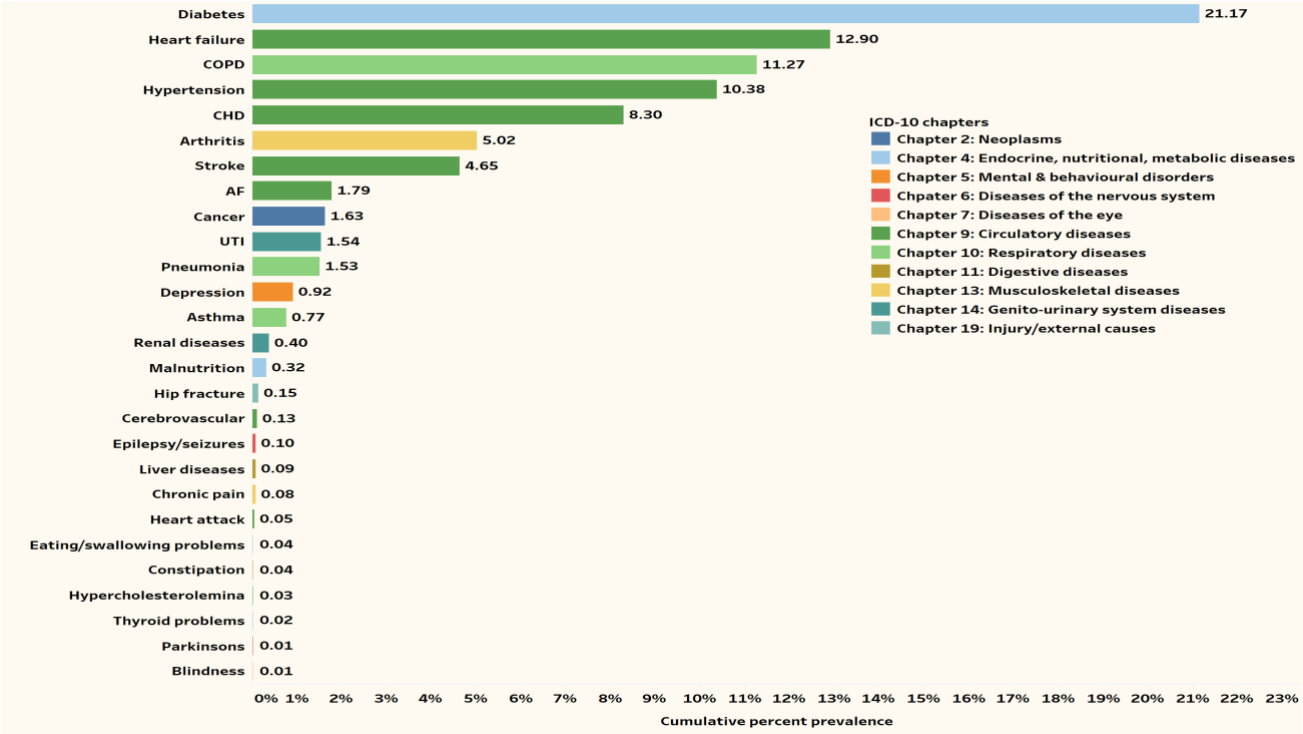


Figure 5: Comorbidities reported in individuals followed in the selected studies.

The numbers represent the weighted percent prevalence. The bar chart shows that diabetes and cardiovascular and cerebrovascular comorbidities were the most prevalent in the selected studies samples. While other non-specific symptoms such as constipation, blindness, chronic pains were less reported probably because they are not coded in the administrative databases and usually as free text which need natural language processing to extract this information.

Abbreviations: COPD: Chronic obstructive pulmonary disease
CHD: Coronary heart diseases
AF: Atrial fibrillation
UTI: Urinary tract infections

2.4 Discussion

I found that the majority of the studies had small sample sizes, largely following only Alzheimer's disease patients, had a short follow-up and few repeat measurements of key outcomes. Only a handful of studies used linked EHRs and there was a paucity of studies incorporating longitudinal measurements of biomarkers for dementia progression.

There was a bias towards North America and Europe and only a few studies originated from Asia. I observed a large disparity between the prevalence of dementia globally compared to the sample size of dementia patients in the selected studies from these regions in our review. This means that dementia progression might be very different in Asia or Africa, but we don't know from the evidence available. The incidence of dementia is increasing at a higher rate in the middle-income countries and therefore, research and studies are desired from these regions.

It is important in longitudinal research studies to know the level of agreement between diagnosis made by different clinicians using the same diagnostic criteria for the same patients at one point in time. I found the revised NINCDS-ADRDA diagnostic criteria were widely used. Studies have found agreement between clinicians using NINCDS-ADRDA criteria varying from "fair" [119] and "moderate" [120] to "substantial" [119,121]. The small number of VaD in the total sample size can be attributed to the fact that only a few studies used subtype-specific diagnostic criteria (such as NINDS-AIREN for VaD) and used ICD codes for sample selection from administrative databases. Finding VaD cases in these databases using ICD codes is challenging since the codes have not kept up with the changes in diagnostic criteria for VaD in the past decades[122]. There is a need to use optimal computational phenotype which consists of a combination of relevant diagnostic codes and codes for medications or tests relevant for the condition of interest. For

example, diagnostic codes consistent with the VaD diagnostic criteria, which include codes for an episode stroke 90 days before the first observed dementia diagnosis[89] to find VaD cases.

The high cost and the invasive procedure of CSF sampling can be one of the reasons behind the lack of studies measuring these biomarkers[123,124] . Apart from the invasive nature of these sampling procedures which can affect participant's compliance with repeated sampling, the radioactive compounds for PET are not available in different geographical locations and therefore, it could be another reason why the majority of the studies did not take repeat measurements of these biomarkers. Therefore, cheaper and less invasive blood biomarkers are needed which are also the focus of research [125]. Apart from their use in diagnosis and enriching trial participants, biomarkers (such as Amyloid PET, Tau Pet, MRI, FDG-PET) also provide information about a drug effect on the underlying disease biology and can help in predicting disease progression and also drug development [126]. Large community-based studies without incorporating biomarkers are important in finding risk factors for dementia however, to understand the extent of their association with dementia will require complementary studies incorporating biomarkers [127]. For disease-modifying therapy (DMT), we will need to design clinical trials that must engage in biological drug targets [36]. DMT will have to show an effect on both biomarkers and clinical symptoms to be approved [128]. Therefore, biomarker data on a subset of community study data will be valuable in understanding the biology of dementia and will also ensure that participants with suspected non-Alzheimer's pathology (SNAP) are not followed in the trials as they will not respond to a DMT.

Progression of dementia is frequently quantified by measuring cognition and it is an important outcome measure in trials as well[113]. I observed that studies were mainly focused on measuring cognition. However, evidence suggests that function and dependency, rather than cognitive decline is a more significant predictor of quality of

life[129]. Therefore, measures such as function, behaviour, and quality of life, which focus on what an individual feel or can do are of greater relevance to patients and carers. The U.S Food and Drug Administration (FDA) also advocates the use of a co-primary outcome measure that incorporates cognitive, functional or a global outcome in drug trials [130]. This approach is required in observational studies involving patients diagnosed with dementia to capture the multidimensional progression of dementia. However, outcome measures of progression will depend on the study goal, population, as at the mild cognitive impairment (MCI) stage, cognition can be the only outcome to measure as the functional impairment might not be evident yet.

Linked EHRs (primary care, secondary care, mortality) contain information about patient's health and medical history and because of their size can offer better statistical power and due to the linkage of different EHRs give a complete picture of patient's clinical care history. We can use the linked EHRs for example, to describe any heterogeneity according to dementia subtypes or co-morbidities and inform clinical trial selection.

The majority of the studies in this review used administrative databases (insurance claims records) as opposed to clinically actionable EHRs and therefore lack the clinical granularity of electronic health records. These studies had larger sample sizes, but the selected samples can be of questionable quality as they only used ICD codes and in most cases without other validation codes such as medications or relevant test codes. These studies were focusing on measuring health system utilisation and did not measure dementia-specific outcomes. Also, the applicability of findings from countries with distinct health and social care system (e.g., Finland, Sweden, Australia, Taiwan) for UK policy is also questionable because of the structure of these different health systems, barriers to accessing healthcare in the different health systems and also the lifestyle of the population and their risk factors. The studies from the UK using EHRs [115,131,132]

(CPRD, THIN) did not use them comprehensively to create clinically-validated trajectories of dementia at a scale in the UK that enables the full heterogeneity of dementia patients experience to be described. These studies had a small sample size (N<5000) and the analysis was restricted to only a few comorbid conditions. A better understanding of variation in dementia progression will provide important insights for improving the treatment, care, and management of dementia patients. The identification of higher resolution dementia patient subgroups with different clinical needs will inform public health planning and help re-structure healthcare delivery for dementia patients. It could also inform the selection of patients for clinical trials of new dementia drugs or other care interventions.

The review takes a multidimensional approach. It captures studies with a range of outcome measures important in dementia progression- from cognitive, functional, Neuropsychiatric to health system utilisation and comorbidities. It covers more than 10 years and includes prospective, retrospective cohort studies to capture as much evidence as possible. The review is only focused on studies with participants diagnosed with dementia aged 65 or over and published in English. Also, the review is based on searching in OVID-MEDLINE alone. However, the search strategy was very comprehensive, covering disease progression, study types, and healthcare utilisation. I might have omitted important studies from the analysis as this was not a full systematic review and more studies can be identified which may have been omitted by our selection criteria.

In summary, I observed a paucity of studies reporting the multidimensional progression of dementia. Studies did not take repeat measurements of key dementia outcomes at several time points over a sufficiently long period and therefore, there are limitations on what can be achieved from these studies about dementia progression. Also, no study has comprehensively used EHR to investigate dementia progression after diagnosis and therefore, the post-diagnosis experience of dementia patients, their interaction with the

healthcare system, co-morbidity profile, and treatment before death remains poorly understood. Furthermore, it remains unknown how this experience may vary according to important sub-groups defined by dementia subtypes or by the co-existence of other comorbidities common in old age.

2.5 Limitations

The review is only focused on studies with participants diagnosed with dementia aged 65 or over and published in English. Also, the review is based on searching in OVID-MEDLINE alone. I also did not assess the quality of the included papers as my aim was characterise what has been studied by the published studies which met my inclusion criteria. I only used OVIDE-MEDLINE database; however, the search strategy was very comprehensive, covering disease progression, study types, and healthcare utilisation. I might have omitted important studies from the analysis as this was not a full systematic review and more studies can be identified which may have been omitted by our selection criteria. Another limitation is that the search criteria was limited to time period util 2019 and an updated search may result in more recent studies.

2.6 Conclusion

It is clear from this review that to better understand dementia progression and inform future observational studies and clinical trials, we need better standardisation in sample selection, longitudinal measurement of key dementia outcomes at several time points in a diverse cohort over a longer time. Also, with the prospect of disease-modifying therapy, we must engage biological targets to understand the temporal dynamics and true aetiology of dementia. The reliability of the results published using insurance databases is a concern and their results may not be generalizable beyond the databases they were derived from.

Using an optimal computational phenotype, linked EHRs longitudinal data can be used to study the heterogeneity of dementia patients based on comorbidity profile or subtype which will improve treatment, care, and management of dementia patients.

Chapter 3: Exploratory data analysis

3.1 Introduction

This chapter describes the data sources used for identifying dementia patients, the structure of these data sources and the selection patients with good quality data to include in the analysis. I also performed some summary statistical analysis and looked at the top comorbidities in dementia patients' records.

3.2 Data sources

The data sources for the initial dementia patient's identification are the clinical practice research datalink (CPRD) and hospital episode statistics (HES). The general description and structure of both of these data sources are described below.

3.2.1 CPRD GOLD

CPRD is a government non-profit research service which supplies anonymized primary care data from the English NHS registered GP practices for health research. CPRD is sponsored jointly by the Medicine and Healthcare products Regulatory Agency (MHRA) and the National Institute of healthcare research (NIHR) which is funded by the Department of Health and Social Care.

CPRD data comes as CPRD Aurum which collect data from UK GP practices using EMIS system. In the May 2022 version of the CPRD Aurum database, the research acceptable patient count stands at 41+ million , with a significant coverage of 19.83% of the UK population, with 13,300,067 patients. A substantial number of patients,

38,377,503, are eligible for linkage. Moreover, the database comprises data from 1,491 general practitioner practices, contributing to approximately 16.45% of the total UK general practices.

I used CPRD GOLD data which is longitudinal routinely collected electronic health records (EHR data) from UK primary care practices using Vision® general practice patient management software. As of June 2022, CPRD GOLD contained 20+ million acceptable patient records from over 985 UK general practices.

The CPRD contains data of 3,069,853 active patients (currently alive and registered) which is approximately 4.58 % and broadly representative of the UK population [133].

The data in my study was from the period between 01/01/1998 and 31/12/2017 as the project ISAC application was accepted and the data was granted in 2017 as the latest batch of available data.

3.2.1.1 CPRD data structure

Data within CPRD is contained within ten different data tables, each containing different types of information. Most of these data tables can be linked together to combine information. Each patient and practice have a unique ID.

A summary of the information contained in each data table is below (Table2).

Table 2: CPRD data tables structure

Data table	Key information	Linkage to other tables	Primarily used for
Practice	• Practice region • Date the practice started reporting research quality data. • Last date the practice reported information	All tables	Defining study population
Patient	• Patient's demographics • Date of (de) registration with practice	All tables except the staff table	Defining the study population and extracting demographic data
Clinical	• Diagnostic codes • Symptoms codes • Date of diagnosis	Patient, additional details, staff, consultation, and practice table.	Extracting information on Diagnoses, symptoms, investigations
Additional details	• Clinical measurement or observations	Clinical, patient and practice	Extracting clinical characteristics data such as weight, BP, smoking
Therapy	• Medications prescribed and dose. • Date of prescription	Patient, staff, consultation, and practice	Extracting information on treatments prescribed
Test	• Test requests • Test results	Patient, staff, consultation, and practice	Extracting clinical and laboratory test data such as cholesterol, BP.
Referral	• Specialty referred to. • Type of referral (inpatient/day visit etc • Date of referral	Patient, staff, consultation, and practice	Extracting data on referrals to secondary care and related diagnoses

Immunization	• Stage of immunization • Methods (oral/intramuscular etc) • Immunization date	Practice, staff, and consultation	Extracting data on immunizations given
Staff	• Staff role and gender	All tables except the patient, additional details tables	Linking clinical information to the staff member recording it
Consultation	• Type, Length, date of consultation • Staff member conducting the consultation.	All tables except additional detail	The context in which clinical information was recorded

3.2.1.2 CPRD data flow

Primary care data are submitted to the CPRD via the software system the participating GP practices use.

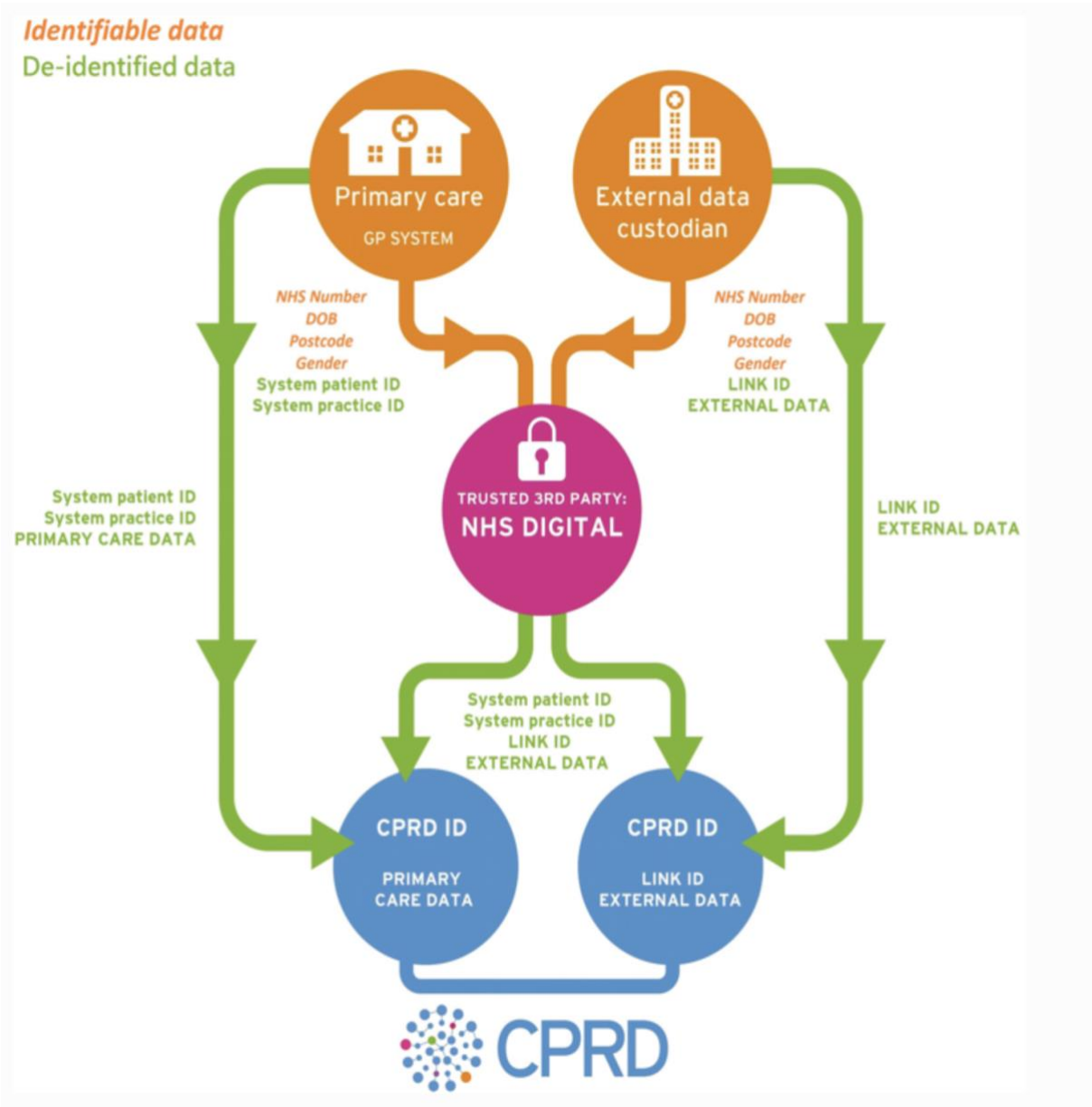


Figure 6: Primary care data flow.

De-identified linked data can either flow from external data custodians to NHS Digital and subsequently to CPRD, or directly from external data custodians to CPRD.
Source: Padmanabhan, et al, 2019 [134]

3.3 Hospital episode statistics (HES)

NHS Digital (formerly HSCIC) administers HES on behalf of the Secretary of State for Health (http://www.digital.nhs.uk/hes. HES contains details of all admissions to NHS hospitals and some admissions to private hospitals, all NHS outpatient appointments, and all A&E attendances in England. HES records include NHS-funded patients treated in English NHS trusts and independent providers and also private patients treated in NHS trusts[135]. In this pilot work, I only used HES inpatient data.

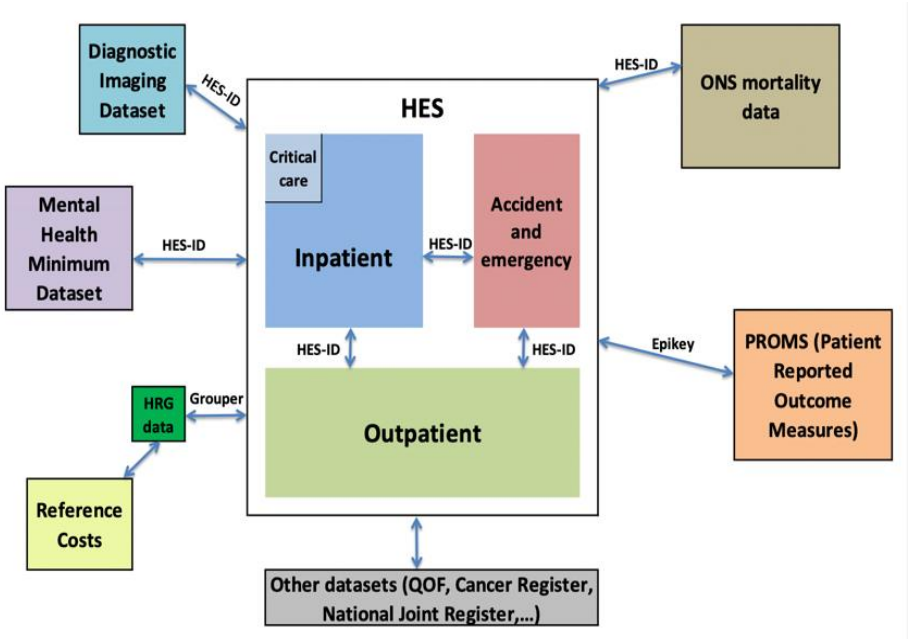


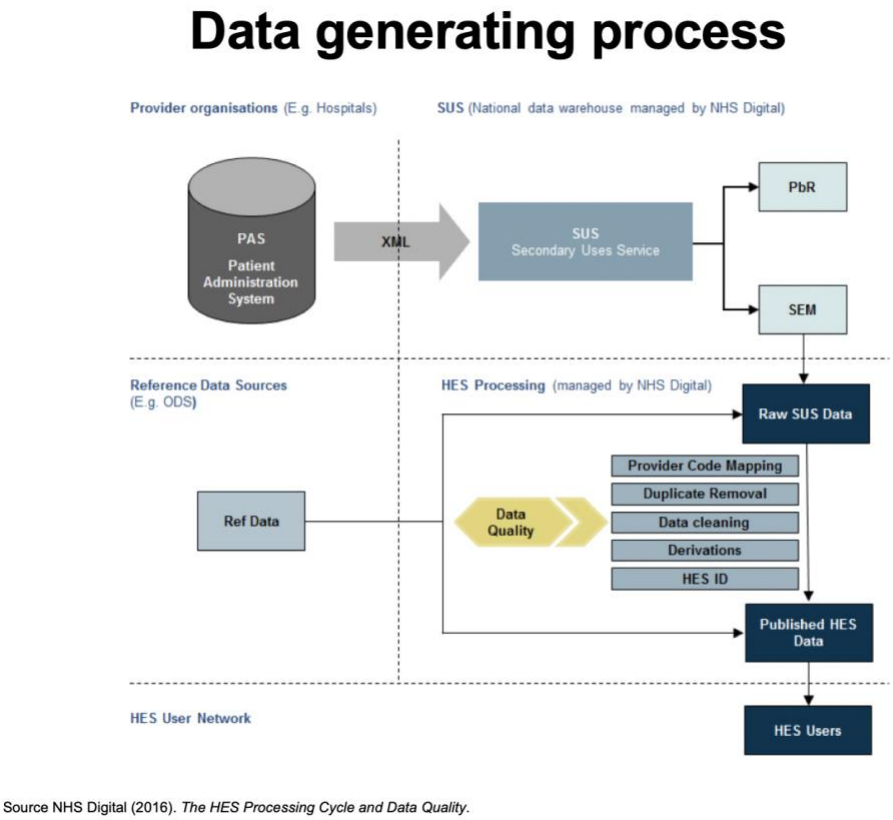
Figure 7: HES data contents and other linkable data

Source: Centre of health economics, University of York

3.3.1 HES data collection

The data is initially recorded during each hospital episode as part of the Commissioning Data Set (CDS) and then made available for secondary service use

(SUS) on a monthly basis[1]. Then data from SUS were extracted for hospital reimbursement under payment by result (pbR) and also a copy of the extract was made available to SUS [136]. After some basic cleaning and checks by the NHS digital and pseudonymized patient IDs are attached to every episode and after the annual refresh for the hospitals the dataset is made available.



Source NHS Digital (2016). *The HES Processing Cycle and Data Quality.*

Figure 8 : HES data generating process.

Source: NHS Digital 2016

3.3.2 HES Data structure

HES admitted patient care data (HES APC) is structured into episodes and spells.

[1] (https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets/commissioning-data-sets, 2020)

A spell is period of continuous care in one provider institution and can have one or more episodes [137]. Each row in the HES APC contains a record of a single finished episode.

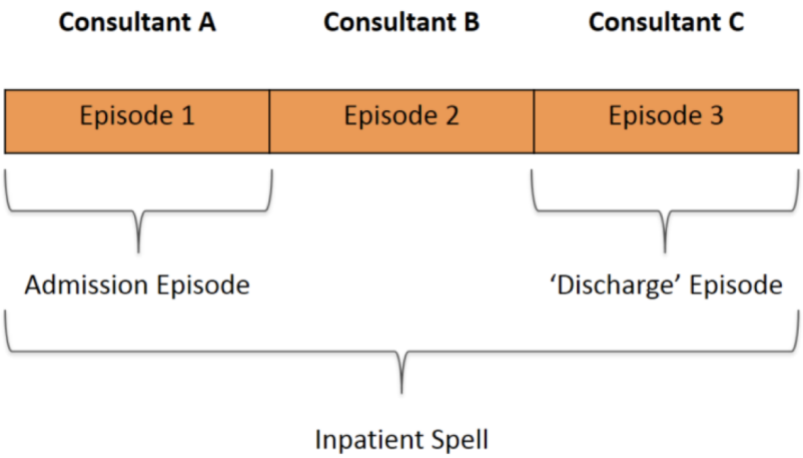


Figure 9: Episodes and spells in HES

Source*: adapted from Methodology to create provider and CIP spells from HES APC data 2014 NHS Digital* [137]

An inpatient spell is a patient's entire stay in one hospital from admission to discharge which may have one or more episodes.

3.3.3 Linkage to other sources

HES APC can be linked to other datasets such as HES A&E, HES outpatients, diagnostic imaging data, mental health, patient reported outcome measure, mortality data using the unique HESID as shown in figure 7 above. HES data can also be linked to clinical practice research data link (CPRD) which is the UK primary care data.

3.4 Methods

3.4.1 Selection of patients records with good quality data.

The study period was selected to be between 01/01/1998 and 31/12/2017. Patients were included if they were aged 30 or over at the start of the study and had been contributing at least one year of up-to-standard (research quality) data. Up to standard data for each patient starts at the latest of the practice data becoming research standard (uts date), patients current registration date (crd) and date of birth. Choosing the latest of these three dates delimits the start of a valid follow up period of the patient.

Similarly, the up-to-standard data for each patient end at the earliest of transfer out date (tod), death date, and last collection date (lcd). Records of patients whose up-to-standard data ends before it starts, were dropped. This happens when a practice up-to-standard date occurs after a patient transfer out of the practice. These people do not contribute any research standard data. The selection of patients eligible for inclusion according to my study start and end date and with research quality data is described in figure 10. The records of these eligible patients will be used for dementia patient's identification and further analyses.

3.4.2 Study entry and exit date for each patient.

I need to know a date on or after 01/01/1998 at which the patient is aged at least 30 and has already been contributing research quality data to the database for at least one year i.e., the latest of study start, start of up to standard data with additional one year and the date at which the patient was at least 30 years old. The study exit date was determined by taking the minimum of study end date and the last date the patient contributed up to standard data. Patients who appear to exit the study before they enter were dropped as well. The maximum follow-up for the eligible patients with research standard data was 19.2 years.

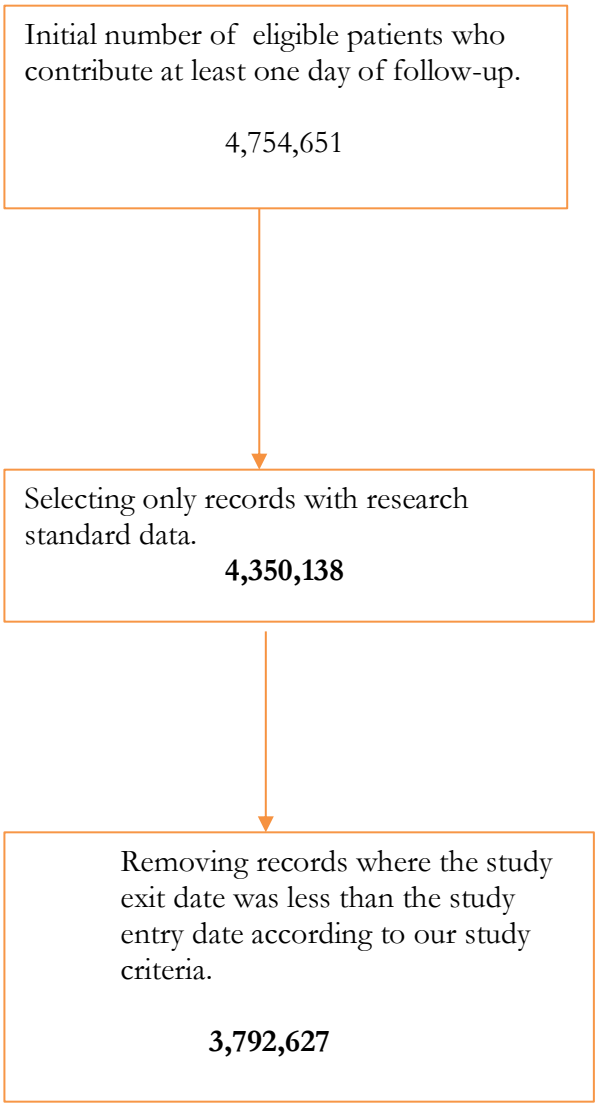


Figure 10: Patient selection process.

Patient's study entry and exit dates were estimated using the method described above. Patients must have at least one day of follow-up in the study. If their estimated study exit time was less than study entry time, then those patients were removed. I removed those records where the study exit date < study entry date.

3.4.2.1 Eligible patients with research standard data

Patients were selected according to the method describe above in section 4.4.2.

Table 3: Total number of eligible patients identified in the CPRD GOLD data linked to HES.

Gender	Number of individuals (%)	Age at entry into the study, mean (SD)
Male	1,871,909 (49%)	46.88 (15.40)
Female	1,920,718 (51%)	49.26(17.52)

Table 3 above shows the number of patients identified to have research quality data and within our study entry and exit date.

Table 4: Diagnosis of dementia in CPRD, HES 1998-2016

CPRD	HES	Frequency	Percentage	Age at entry (mean, (SD)	Gender Male (%)	Follow Up Mean, (SD)
No	No	3,612,510	95.3%	46.76 (15.59)	50%	6.87 (5.63)
Yes	No	16,225	0.4%	77.58 (11.4)	29.8%	6.26 (5.96)
No	Yes	119,813	3.2%	73.80 (11.71)	38.9%	6.81 (5.23)
Yes	Yes	44,079	1.2%	77.50 (9.44)	33%	6.25 (5.40)

Table 4 above shows the number of patients identified to have a record of a dementia diagnosis code in each database, their gender percentage, mean age and follow-up.

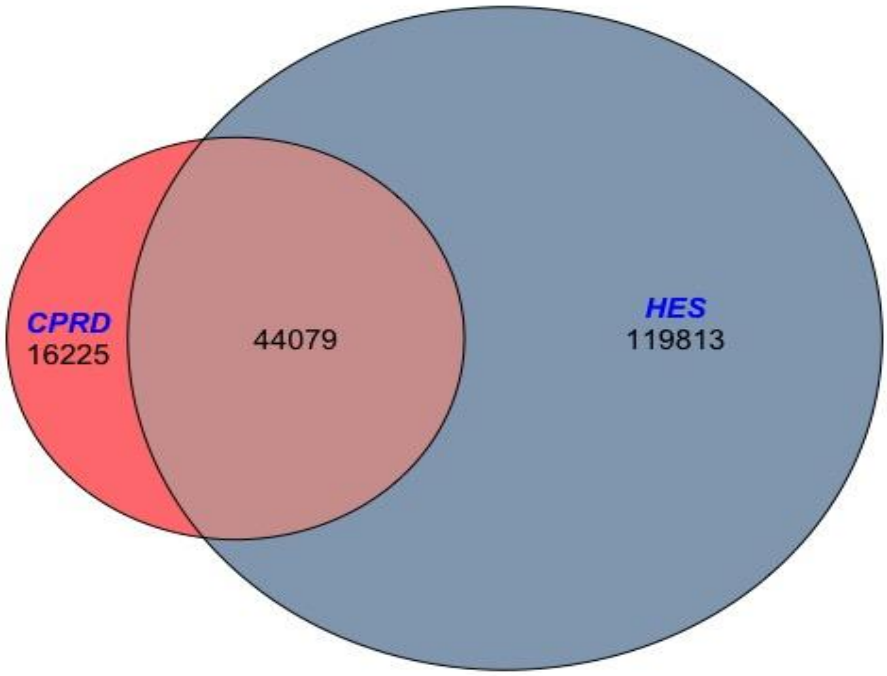


Figure 11: Number of dementia patients identified in CPRD and HES.

CPRD = Clinical Practice Research Datalink

HES = Hospital Episode Statistics

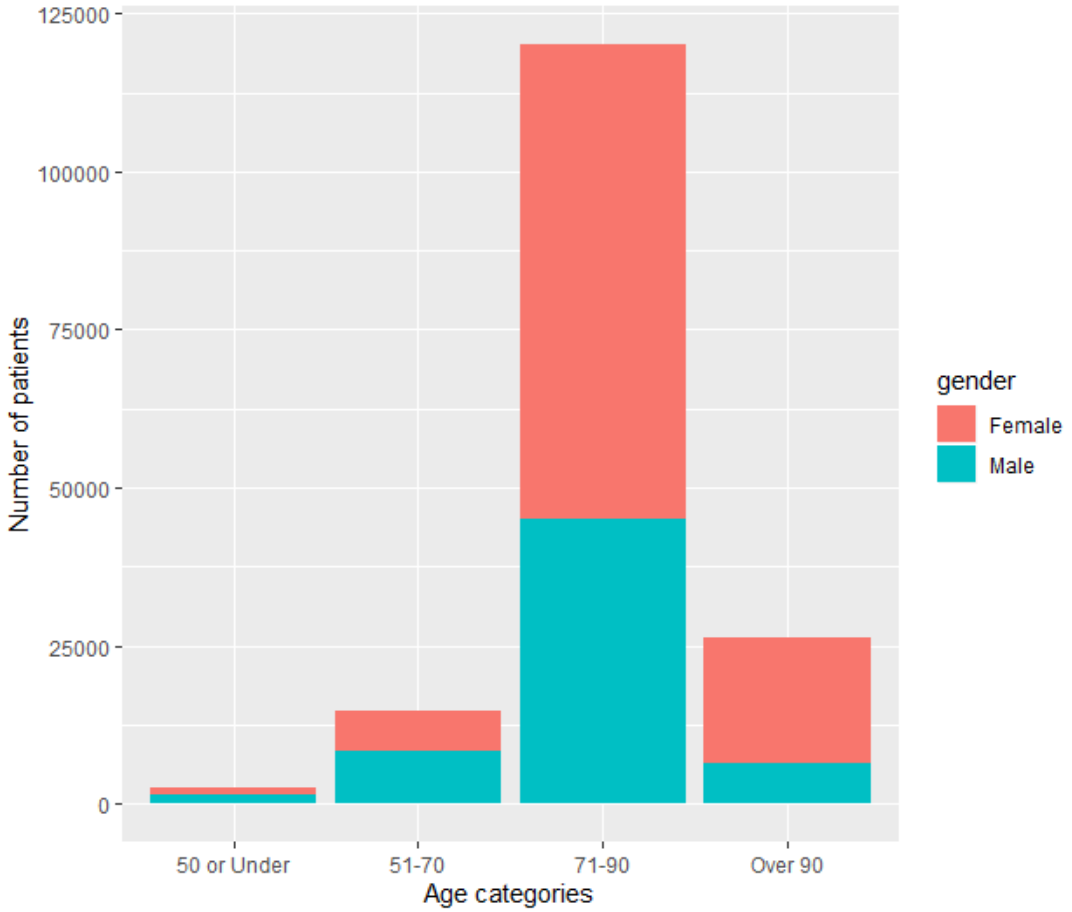


Figure 12: Dementia patients by age groups and gender.

The bar chart shows that majority of dementia patients were in the age group 71-90 and of female gender.

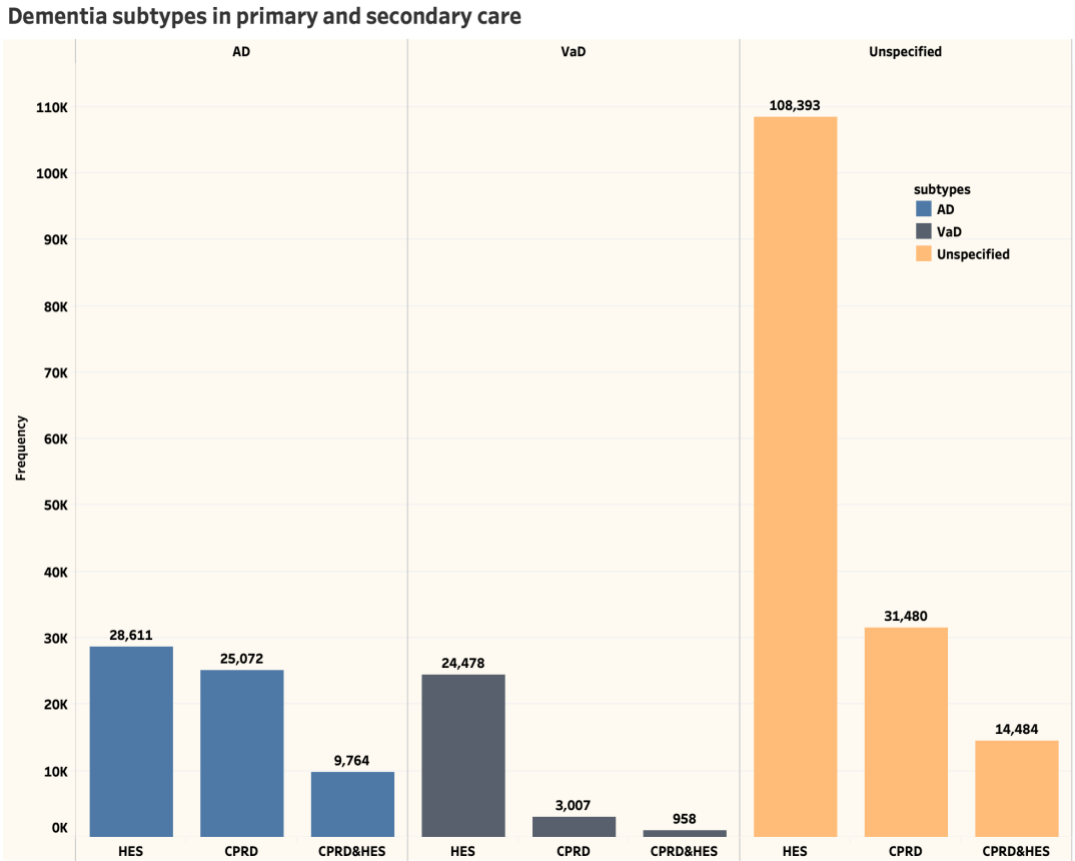


Figure 13: Dementia subtypes in HES and CPRD

The number of patients with diagnostic codes for different dementia subtypes in primary and secondary data. Majority of codes were for unspecified dementia.

AD = Alzheimer's dementia

VaD = Vascular dementia

CPRD = Clinical Practice Research Datalink

HES = Hospital Episode Statistics

Figure 13 shows the dementia subtypes frequency in each CPRD, HES alone and combined. We can see that there is a big difference in the number of VaD cases between the CPRD and HES databases as majority of VaD cases were from HES. Diagnosing VaD cases in the primary care using only read codes is challenging since the codes have

not kept up with the changes in diagnostic criteria for VaD in the past decades and it could also be due to the lack of specialist knowledge and tests available at the primary care clinics. Also the low diagnostic percentage could be due to the fact that VaD account of only 20% of dementia cases and it also require additional criteria for conformation of diagnosis such as using biomarkers data, imaging and making sure that the diagnostic codes consistent with the VaD diagnostic criteria, which include codes for an episode stroke 90 days before the first observed dementia diagnosis[89] to find VaD cases.

Table 5 shows the top 30 most prevalent ICD codes for dementia patients one year before their dementia diagnosis. The percentages represent the number of unique patients with the ICD-10 codes for the associated conditions within our dementia cohort.

Table 5: The top 30 ICD codes in HES for dementia patients before their diagnosis

Number	ICD	Description	Frequency	Percentage
1	I10	Essential (primary) hypertension	284474	6.5
2	E119	Type 2 diabetes mellitus without complications	114941	2.6
3	I48	Atrial fibrillation and flutter	109945	2.5
4	Z867	Personal history of disease of circulatory system	73042	1.7
5	N390	Urinary tract infection, site not specified	64529	1.5
6	I209	Angina pectoris, unspecified	63956	1.5
7	E780	Hypercholesterolaemia	55938	1.3
8	J459	Asthma, Unspecifed	49525	1.1
9	I259	Chronic ischemic heart disease, unspecified	48863	1.1
10	Z921	Personal history of long term use of anti-coagulant	46173	1
11	E039	Hypothyroidism, unspecified	42016	1
12	H269	Unspecified cataract	40804	0.9
13	Z864	Personal history of psychoactive substance abuse	38980	0.9
14	I258	other form of chronic ischemic heart disease	37498	0.9
15	J449	Chronic obstructive pulmonary disease, unspecified	36474	0.8
16	Z491	Renal dialysis	35767	0.8
17	Z966	presence of orthopaedic joint implants	34845	0.8
18	R410	Disorientation, unspecified	30073	0.7
19	D649	Anemia, unspecified	29934	0.7
20	Z880	Allergy status to penicillin	29502	0.7
21	K449	Diaphragmatic hernia without obstruction or gangrene	29167	0.7
22	Z922	Long term use of other medicaments	28445	0.6
23	M199	Arthrosis of first carpometacarpal joint, unspecified	27884	0.6
24	Z602	Problems related to living alone	27851	0.6
25	I500	Congestive heart failure	27123	0.6
26	I251	Atherosclerotic heart disease	26529	0.6
27	F329	Major depressive disorder, single episode, unspecified	26485	0.6
28	R55	Syncope and collapse	26281	0.6
29	Z951	Presence of aortocoronary bypass graft	25103	0.6
30	R69	Illness, unspecified	25035	0.6

3.5 Summary

In this exploratory analysis I found that out of the total patients with good quality data (3,792,62), 0.4% (16225) had dementia codes alone in CPRD, 3.2% (119,813) in HES and 1.2% (44,079) in both CPRD and HES. The percentage of male dementia identified in both CPRD, and HES was 33% and the age at entry into the study was 77.5 years (table 4). Majority of dementia patients were in the age group 71-90 and unspecified dementia was the most common recorded diagnosis in both CPRD and HES, followed by AD and VaD. Diabetes, cardiovascular complications were the most occurring comorbidities in the dementia patients. I will explore the association of diabetes with dementia diagnosis and other associated factors in the next chapter.

Chapter 4: Time to event analysis of dementia diagnosis in diabetic patients

4.1 Introduction

Type 2 diabetes (T2D) and dementia are two major health crises which are occurring simultaneously, and evidence suggests that T2D and related features such as insulin resistance are associated with increased dementia risk [138,139]. In my literature review on dementia progression studies and preliminary analysis of HES data, diabetes was the top comorbidity.

4.2 Objective

The objective was to identify the factors in patients with diabetes associated with the risk of developing dementia.

4.3 Methods

4.3.1 Study design, data source and patient selection

The study is a retrospective longitudinal study using the hospital episode statistics (HES) data for patients' identification and their records for anti-diabetic medications from the clinical practice research data (CPRD). Patients who were diagnosed with diabetes between 01/01/1998 and 31/12/2007 were included. Patients who were diagnosed with dementia or died before our index date of 01/01/2008 were excluded. Only people who were alive and were not diagnosed with dementia by 31/12/2007 were followed in the study up to the study end date of 31/12/2017. The patient selection process is shown in figure 14. A total of 68,055 patients met the inclusion and exclusion criteria. HES admitted patient care (HES APC) records were used to identify patients with a diagnosis of diabetes in their records using the ICD 10 codes for diabetes. ICD codes used for

diabetes identification were E10, E11, E12, O242, E13, E14, G590, G632, H280, H360, M142, N083, O240, O241, and O243.

Anti-diabetic medication codes were identified by using the CPRD code browser to identify product codes for anti-diabetic medications. In addition to Insulin the different classes of anti-diabetic medications (ADD) were biguanides (metformin), sulfonylureas such as glibenclamide, gliclazide, glimepiride, gliplizide, glinides (repaglidine), glitazones (pioglitazone, rosiglitazone) and dipeptidyl peptidase-4 inhibitors such as saxagliptin, sitagliptin and vildagliptins. The CPRD GOLD therapy file was then used to extract information about the use of ADD for the patients identified with diabetes.

4.3.2 Patients' characteristics

The sociodemographic characteristic of the included cohort is grouped into those with and without dementia is shown in Table 6. The median age of the total cohort was 70 and the cohort was comprised of ~ 54% men and 89% of the cohort was of white race.

The cumulative incidence of dementia between January 01, 2008, and December 31, 2017, was 13.8% with majority of the dementia subtypes were coded as unspecified dementia (50.3%) while vascular dementia was (19.4%) and Alzheimer's dementia ICD codes were present in 11.7% of the cohort included in the analysis. People who developed dementia were mostly white (92.4%) and in the age groups of 75-84.

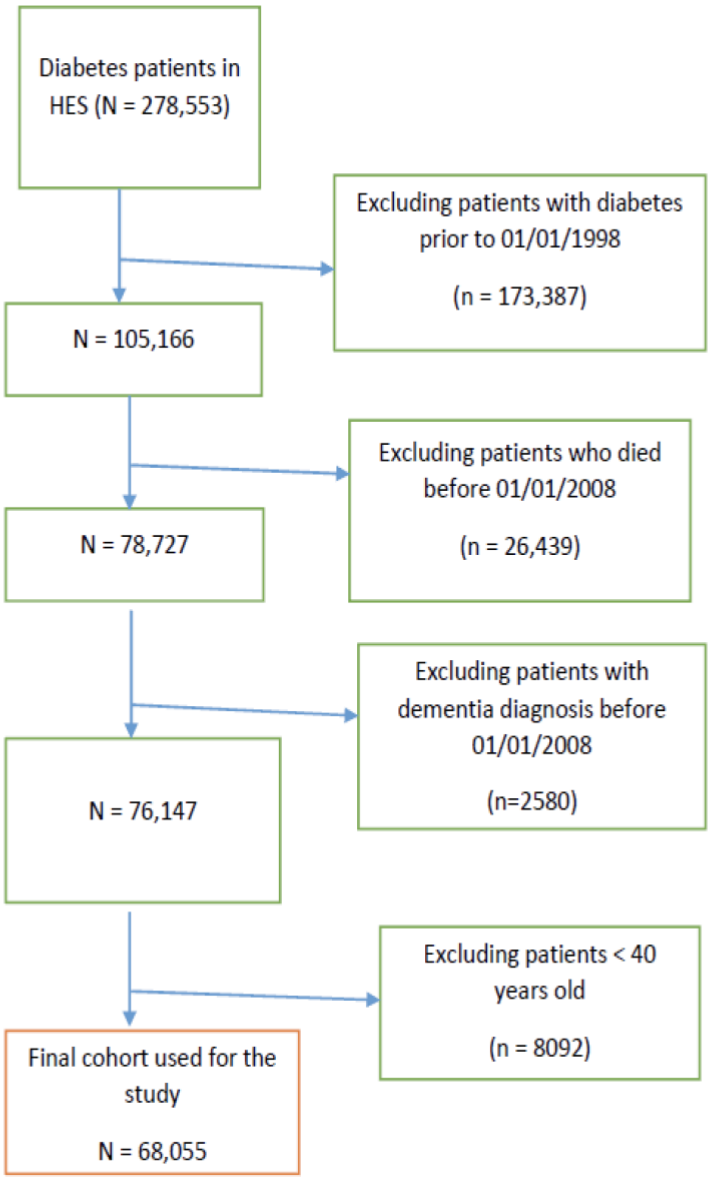


Figure 14: Selection of study sample.

4.3.3 Covariates

Age is known to be associated with dementia and therefore, patients' ages at the start of follow–up (index date of 01/01/2008) were used and categorised into under 65, 65-69, 70-74, 75-79, 80-84, 85 and over. Ethnicity as recorded in hospital, was grouped into white, Asian, Black, and mixed/other. Marital status was grouped into married, unmarried, unknown and widowed/separated. Patient practice region information was also used in the model which group England into 11 different regions. I used this information along with the index of multiple deprivation (IMD), which was grouped into quintiles to understand the effect of socio-economic factors and the effect of patient care for diabetes in different regions of England.

4.3.4 Exposure

Anti-diabetic drug exposure is the use of these medications during the year 2007. Patients were classified as Y/N for exposure to insulin and oral anti diabetic drugs mentioned above. The duration of diabetes was grouped into three categories: patients diagnosed with diabetes between 01/01/2006 and 31/12/2007 were considered to be diabetic for 1-2 years by our study index date of 01/01/2008. Similarly, patients diagnosed between 01/01/2003 and 01/01/2006 were considered to be diabetic for 3-5 years while those diagnosed between 01/01/1998 and 01/01/2003 for 6 or more years. The model was also adjusted for other comorbidities such as cerebrovascular complications, obesity and also for incident of hip fracture.

4.4 Outcome

4.4.1 Time to dementia diagnosis

Dementia was defined as the first occurrence of the relevant ICD codes for dementia from the index date of 01/01/2008. Patients were followed up for the occurrence of first diagnosis of dementia. The ICD codes used for dementia identification are attached [Appendix A].

4.5 Statistical analysis

A Cox proportional hazards model was used to calculate the hazard ratios (HRs) for the effect of each predictor variable upon the occurrence of dementia after adjusting for potential covariates. Patients were censored at the end date of the study (31/12/2017) or if they died before the diagnosis of dementia. The overall crude incidence of dementia diagnosis in diabetic patients over 40 between 2008 and 2017 was 18.9 cases per 1000-person years and was calculated by dividing the total number of incident dementia cases by the total time at risk for all the patients in the study. All data management and statistical analyses were conducted using R 3.6 software [140].

4.6 Results

The Cox proportional hazard model was used to investigate the independent association of different variables with the incidence of dementia in diabetic patients. Table 7 shows the results from the Cox regression analysis. Men have a lower risk (HR, 0.87: CI 0.84-0.91: p <0.001) and as expected old age was also a high risk for dementia. People in the age group 75-79 years had almost 10 times higher risk (HR, 9.69: CI 8.93- 10.51, p <0.001) compared to people under 65 years old. Similarly, people in the age group 80-84 had almost 16 times higher risk (HR, 15.70; CI 14.46-17.05, p <0.001) and in the age group 85 and over more than 23 times higher risk (HR, 23.59; CI 21.65 – 25.70, p <0.001)

compared to people under 65 years old. This means that at any time in the follow-up dementia occurred in 23 times as many patients in the "85 & over" group with diabetes, proportionally to the under 65 years old group.

The risk of dementia was lowered in the Asian ethnic minorities (HR, 0.64; CI 0.57 – 0.72, p <0.001) compared to the white race while black ethnic minorities risk was lowered compared to white as well however, it was not statistically significant (HR, 0.98, CI 0.85-1.12, p = 0.732).

People with diabetes who were widowed or separated had a significantly higher risk of dementia (HR, 1.21; CI 1.09-1.35, p <0.001). The risk of dementia was significantly higher in the greater London (HR, 1.19; CI 1.09 – 1.29, p <0.001), East Midland (HR, 1.15; CI 1.03-1.29, p 0.016), West Midland (HR ,1.14; CI 1.04-1.24, p 0.004) and Northwest of England (HR 1.16; CI 1.08 – 1.26, p <0.001) regions [Figure 15]. Also, people living in the areas which were among the 20% most deprived area (IMD quintile 4) were at significantly higher risk (HR 1.18; CI 1.11-1.25; p<0.001) than people in the least deprived areas (IMD quintile 1).

The risk of dementia was significantly less (HR, 0.89; CI 0.86-0.93; p <0.001) in patients who had prior use of insulin or other anti-diabetic (ADD) medication (biguanides, sulfonylureas, glitazones). The duration of diabetes was also significantly associated with the development of dementia. With the 1-2 years of diabetes duration as the referent, the HR of diabetes of the 3-5 years and 6 or more years were (HR, 1.14; CI 1.08-1.20; p <0.001) and (HR, 1.24; CI 1.18-1.31, p <0.001) respectively. The presence of cerebrovascular complications and incidence of hip fracture were also significantly associated with the risk of dementia with a HR (95% CI) of 1.57 (1.43-1.71) and 1.51 (1.26 – 1.81). Obesity was not statistically significant with the risk of dementia.

Table 6 : Comparison of baseline characteristics of patients with vs patients without dementia

	Dementia (n=9424)	No dementia (n=58631)	Total (n=68055)
Gender			
Female	5208 (55.3%)	26327 (44.9%)	31535 (46.3%)
Male	4216 (44.7%)	32304 (55.1%)	36520 (53.7%)
Age			
Mean (SD)	77.4 (8.73)	68.1 (11.9)	69.4 (11.9)
Age groups			
Under 65 yrs	792 (8.4%)	22804 (38.9%)	23596 (34.7%)
65-69 yrs	800 (8.5%)	8126 (13.9%)	8926 (13.1%)
70-74 yrs	1488 (15.8%)	8874 (15.1%)	10362 (15.2%)
75-79 yrs	2206 (23.4%)	8002 (13.6%)	10208 (15.0%)
80-84 yrs	2216 (23.5%)	5798 (9.9%)	8014 (11.8%)
85 & over	1922 (20.4%)	5027 (8.6%)	6949 (10.2%)
Marital status			
Married	1246 (13.2%)	8679 (14.8%)	9925 (14.6%)
Unknown	7609 (80.7%)	47158 (80.4%)	54767 (80.5%)
Unmarried	107 (1.1%)	927 (1.6%)	1034 (1.5%)
Widowed/separated	462 (4.9%)	1867 (3.2%)	2329 (3.4%)

	Dementia (n=9424)	No dementia (n=58631)	Total (n=68055)
Ethnicities			
White	8708 (92.4%)	51831 (88.4%)	60539 (89.0%)
Asian	316 (3.4%)	3510 (6.0%)	3826 (5.6%)
Black	236 (2.5%)	1342 (2.3%)	1578 (2.3%)
Others	164 (1.7%)	1948 (3.3%)	2112 (3.1%)
Dementia subtypes			
Alzheimer's disease	1101 (11.7%)	0 (0%)	1101 (1.6%)
Possible dementia	1652 (17.5%)	--	1652 (2.4%)
Rare dementia	102 (1.1%)	--	102 (0.1%)
Unspecified dementia type	4743 (50.3%)	--	4743 (7.0%)
Vascular dementia	1826 (19.4%)	--	1826 (2.7%)
England's regions			
Southeast Coast	1221 (13.0%)	7467 (12.7%)	8688 (12.8%)
East Midlands	398 (4.2%)	2332 (4.0%)	2730 (4.0%)
East of England	1053 (11.2%)	6838 (11.7%)	7891 (11.6%)
London	1340 (14.2%)	7969 (13.6%)	9309 (13.7%)
Northeast	208 (2.2%)	1397 (2.4%)	1605 (2.4%)
Northwest	1642 (17.4%)	9643 (16.4%)	11285 (16.6%)
South Central	908 (9.6%)	6049 (10.3%)	6957 (10.2%)
Southwest	1229 (13.0%)	7890 (13.5%)	9119 (13.4%)
West Midlands	1036 (11.0%)	6304 (10.8%)	7340 (10.8%)

	Dementia (n=9424)	No dementia (n=58631)	Total (n=68055)
Yorkshire & the Humber	389 (4.1%)	2742 (4.7%)	3131 (4.6%)
Index of Multiple Deprivation			
1(least deprived	2153 (22.8%)	12901 (22.0%)	15054 (22.1%)
2	2495 (26.5%)	15768 (26.9%)	18263 (26.8%)
3	2375 (25.2%)	14992 (25.6%)	17367 (25.5%)
4(most deprived)	2395 (25.4%)	14891 (25.4%)	17286 (25.4%)
Missing	6 (0.1%)	79 (0.1%)	85 (0.1%)
Diabetes duration			
1-2 years	2610 (27.7%)	18736 (32.0%)	21346 (31.4%)
3-5 years	3316 (35.2%)	20495 (35.0%)	23811 (35.0%)
6 or more years	3498 (37.1%)	19400 (33.1%)	22898 (33.6%)
Prior Insulin/ADD use			
	5228 (55.5%)	35355 (60.3%)	40583 (59.6%)
Cerebrovascular			
	504 (5.3%)	2464 (4.2%)	2968 (4.4%)
obesity			
	97 (1.0%)	1017 (1.7%)	1114 (1.6%)
Hip fracture			
	121 (1.3%)	353 (0.6%)	474 (0.7%)

Table 7: Risk factors for dementia in diabetic patients: results from the Cox proportional hazard regression model

Characteristic	HR[1]	95% CI[1]	p-value	PH Test
Gender				0.075
Female	—	—		
Male	0.89	0.86, 0.93	<0.001	
Age	1.11	1.10, 1.11	<0.001	0.083
Ethnicities				0.062
White	—	—		
Asian	0.65	0.58, 0.73	<0.001	
Black	1.02	0.89, 1.17	0.8	
Others	0.59	0.51, 0.69	<0.001	
England Regions				0.4
East of England	—	—		
Southeast Coast	1.05	0.97, 1.14	0.2	
East Midlands	1.16	1.03, 1.30	0.011	
London	1.20	1.11, 1.31	<0.001	
Northeast	1.05	0.90, 1.22	0.5	
Northwest	1.17	1.08, 1.27	<0.001	
South Central	1.04	0.95, 1.14	0.4	
Southwest	0.99	0.91, 1.07	0.7	
West Midlands	1.14	1.05, 1.24	0.003	
Yorkshire & the Humber	0.97	0.86, 1.09	0.6	
Index of Multiple deprivation				>0.9
1(least deprived	—	—		
2	1.02	0.96, 1.08	0.6	
3	1.12	1.05, 1.18	<0.001	
4(most deprived)	1.20	1.13, 1.27	<0.001	

Characteristic	HR[1]	95% CI[1]	p-value	PH Test
Missing	0.53	0.24, 1.18	0.12	
Diabetes Duration				0.2
1-2 years	—	—		
3-5 years	1.12	1.07, 1.18	<0.001	
6 or more years	1.21	1.15, 1.27	<0.001	
Cerebrovascular				0.4
No	—	—		
Yes	1.58	1.44, 1.73	<0.001	
Obesity				0.6
No	—	—		
Yes	1.16	0.95, 1.43	0.14	
Hip fracture				0.5
No	—	—		
Yes	1.38	1.15, 1.65	<0.001	

[1]HR = Hazard Ratio, CI = Confidence Interval

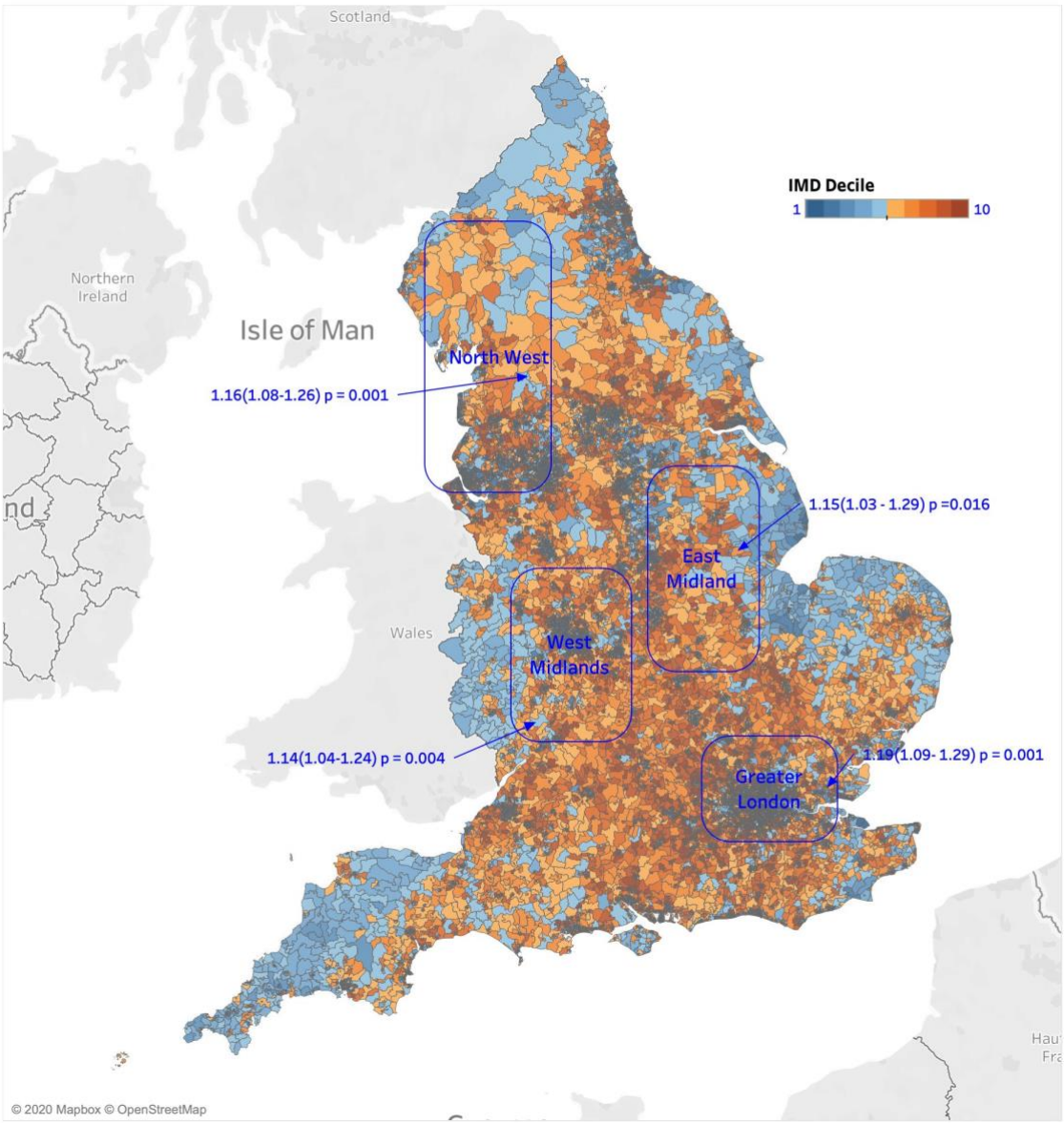


Figure 15: England's regions with statistically significant hazard ratios for dementia.

Figure 15 shows the regions in England which were statistically significantly associated with the risk of dementia. The map also shows the areas of multiple deprivation (IMD) deciles, and these regions contain the most deprived local authorities (orange/red colours). The figure shows England's regions in terms of their socioeconomic deprivation and the regions coloured as red and dark red have areas which are in the most deprived IMD levels, while the blue regions are least deprived. The numbers with 95% CI and p values are the HR from the Cox regression model which shows the HR of these regions for dementia diagnosis. As we can see these are the areas which are in most deprived socioeconomic regions, and they are also the regions which are the diabetic hotspots and high concentration of Black and Ethnic minorities population.

4.6.1 Proportional Hazard Assumption

The proportional hazard assumption was checked by including the hypothesis test (PH Test column) in the model result table, which signifies whether the gradient differs from zero for each variable. It is evident that no variable significantly differs from zero at the 5% significance level because the PH test values are greater than 0.05.

4.6.2 Competing Risk Regression

I also used a competing risk regression using the Fine-Gray model which gives us the instantaneous rate of occurrence of the event of interest in subjects who have not yet experienced an event of that type. Here, I considered death as a competing risk to dementia diagnosis because patients can die before they are diagnosed with dementia and preclude the occurrence of our event of interest. All other variables are kept the same as the Cox proportional hazard model. I used competing risk regression to check what are the factors which affect dementia diagnosis in diabetic patients. I used diabetic duration

as a covariate and adjusted for the effect of age, sex, ethnicity, socioeconomic deprivation, regions of england where the patients were getting care and their comorbidities such as cerebrovascular, obesity and having the event of hip fracture.

Table 8 : Risk factors for dementia in diabetic patients: results from the Competing risk regression model (Fine-Gray Model)

Characteristic	HR[1]	95% CI[1]	p-value
gender			
Female	—	—	
Male	0.83	0.80, 0.87	**<0.001**
Age	1.07	1.07, 1.07	**<0.001**
Ethnicities			
White	—	—	
Asian	0.75	0.66, 0.84	**<0.001**
Black	1.22	1.07, 1.39	**0.003**
Others	0.54	0.46, 0.63	**<0.001**
England Regions			
East Midlands	—	—	
East of England	0.89	0.79, 1.01	0.061
London	1.11	0.99, 1.25	0.068
Northeast	0.91	0.77, 1.08	0.3
Northwest	1.03	0.92, 1.15	0.6
South Central	0.90	0.80, 1.01	0.084
Southeast Coast	0.95	0.85, 1.07	0.4
Southwest	0.87	0.77, 0.97	**0.017**
West Midlands	0.96	0.85, 1.08	0.5
Yorkshire & the Humber	0.83	0.72, 0.96	**0.011**
Index of Multiple deprivation			
1(least deprived	—	—	

Characteristic	HR[1]	95% CI[1]	p-value
2	0.98	0.92, 1.04	0.4
3	1.05	0.98, 1.11	0.2
4(most deprived)	1.10	1.04, 1.17	**0.002**
Missing	0.44	0.20, 0.99	**0.048**
Diabetes Duration			
1-2 years	—	—	
3-5 years	1.10	1.04, 1.15	**<0.001**
6 or more years	1.14	1.08, 1.20	**<0.001**
Cerebrovascular			
No	—	—	
Yes	1.07	0.98, 1.18	0.2
Obesity			
No	—	—	
Yes	0.96	0.79, 1.18	0.7
Hip fracture			
No	—	—	
Yes	1.08	0.88, 1.31	0.5

[1]HR = Hazard Ratio, CI = Confidence Interval

In the Cox proportional hazards model, the hazard ratio (HR) for dementia diagnosis for black ethnic group, compared to white individuals was 1.02, with a 95% confidence interval (CI) ranging from 0.89 to 1.17. The p-value was 0.8, indicating that there is no statistically significant difference in the risk of dementia diagnosis between black and white ethnic groups when death is treated as a censoring event.

In the competing risks model, the sub-distribution hazard ratio (SHR) for dementia diagnosis for black ethnic groups compared to white ethnic group is 1.22, with a 95% confidence interval ranging from 1.07 to 1.39. The p-value is 0.003, indicating that there is a statistically significant higher risk of dementia diagnosis for black individuals compared to white individuals when accounting for the competing risk of death. The Fine-Gray model specifically accounts for the presence of competing risks (in this case, death). It provides sub-distribution hazard ratios that describe the effect of covariates on the cumulative incidence function of the event of interest, accounting for the fact that individuals who experience the competing event are no longer at risk for the event of interest. The Competing Risks Model indicates that black ethnic individuals have a significantly higher risk of being diagnosed with dementia compared to the white individuals when considering that some individuals may die before being diagnosed. This suggests that when accounting for the competing risk of death, the risk of dementia diagnosis is actually higher for the black ethnic groups. The results from the competing risk regression are slightly different with lower HR as shown in table 8 above.

The competing risks model provides a more nuanced and accurate understanding of the risk of dementia diagnosis in the presence of the competing risk of death. The significant result in the competing risks model highlights the importance of considering competing events in survival analysis, especially when those events can affect the occurrence of the primary event of interest.

4.7 Discussion

In a time to event analysis, during a 10-year interval (01/01/2008 - 31/12/2017), dementia developed in 13.85% of patients over the age of 40 who had diabetes from 1-2 years to 6 or more years. This was an exploratory study to better understand how diabetes-related factors and sociodemographic factors in diabetic patients are associated with dementia.

Increased risk of dementia was associated with increasing age, female gender, white race, residence in the 3rd and 4th quintiles of the most deprived 20% of small areas of England. Diabetic patients who were widowed or separated were at higher risk of dementia. Also, patients registered with practices in the Midlands, London area and the Northwest of England were also at statistically significant risk of dementia.

Among the clinical factors associated with the increased risk of dementia were the duration of diabetes while the use of insulin and anti-diabetic medications were associated with a decreased incidence of dementia. Similarly, exposure to cerebrovascular diseases or hip fracture had statistically significant associated risk of dementia.

Both dementia and diabetes are among the most prevalent diseases in the older population [141]. Type 2 diabetes (T2D) is the most common diabetes type and it affects 422 million people globally [142]. Many studies including the Rotterdam study shows and support the elevated risk of dementia associated with T2D [138,139,143,144]. Both diabetes and dementia are long term conditions and it is estimated that by 2025, the number people with dementia in the UK will rise over 1 million and people with diabetes are predicted to be over 5 million [145,146].

I found that 13.85% of diabetic patients have dementia diagnosis codes over the 10 years follow-up period with an incident rate of 18.9 cases per 1000-person years of follow-up. A population-based study also found a 16% increased risk of dementia in newly diagnosed diabetic patients aged over 65 [147].

It has been established that age is an important risk factor for dementia [148] and I found in this study that the risk of dementia increased with age as explained in the results.

The incidence of dementia was 36% lower in the Asian ethnic group compared to the white group. Asian people are exposed to several life style, genetic and environmental risk factors for dementia such as cardiovascular, lower formal education when compared to white group [149]. However, different cultural and genetic factors may contribute to the difference in susceptibility to dementia [149]. Another factor for the lower incidence in the Asian community can be attributed to the reluctance of getting a diagnosis because of the stigma attached with dementia or the fear of moving into a care home [150].

People who had lost a partner through death or separation had a 21% higher risk of dementia compared to married people. This association is similar to that reported by a meta-analysis [151], which reported a 20% increased risk of dementia this group. Patients who were getting treatment for diabetes in the London, Midlands, and Northwest of England, were at statistically significant higher risk of dementia. These are the regions which are the most ethnically diverse with London having 40% of its residents belonging to Black, Asian and minority ethnic groups (BAME), while West Midland, East Midland, and North West regions have 17.4%, 11% and 10% respectively, according to the England and Wales 2011 census [152]. Also, these regions are in the diabetic hotspots: in 2018/19 the Midlands had 7.6% of people on the NHS register as diabetic which is the highest in England, the North West had a 7.2%, the third highest in England [153,154]. Similarly, residents in the areas which were in the 20% most deprived 3rd or 4th quintile (most deprived) compared to quintile 1(least deprived) according to the index of multiple

deprivation 2015, were at higher risk of dementia. Also, the above geographical regions have the most deprived local authorities compared to other parts of England [155].

I found that an increased diabetes duration was associated with increase in dementia risk. The risk of dementia increased from 14% in patients with diabetes for 3-5 years to 24% in patients with diabetes for 6 or more years. Similar results were found by Bruce et al, who found that dementia risk increased by 45% for every 5 years increase in diabetes duration [156]. These finding suggest that duration of diabetes is an important risk factor for dementia.

I found the use of insulin and anti-diabetic medications significantly reduced the risk of dementia. This results support those obtained by Beeri et al, who examined the brain autopsies of diabetic and non-diabetic patients matched on age and sex and found that individuals who had history of combined insulin and anti-diabetic medications had significantly less amyloid plaques which suggests a possible beneficial effect of ADD medication on AD pathology [157]. Similarly, Hsu et al, conducted a study using data from Taiwan's National Insurance database of individuals aged 50 and over with both diabetic and non-diabetic patients and found more than 2-fold increase in the risk of dementia in individuals with type 2 diabetes. They also found that metformin use reduced the risk of dementia by 24% compared to no-metformin use [158]. I found that cerebrovascular was associated with a 57% increase in risk of dementia. Similar results are shown in a population-based study where they found a 2-fold increase in dementia by cerebrovascular conditions [147].

4.8 Strengths and limitations

The study has certain strengths such as the use of hospital episode statistics (HES) for diagnosis identification and CPRD for information on exposure to anti-diabetic medication. HES diagnoses are clinically validated. Secondly, the study has a large cohort size and follow up duration of 10 years. The data used is geographically diverse and covers different regions of England.

As this study was an exploratory study to better understand how diabetes-related factors and sociodemographic factors in diabetic patients are associated with dementia. I was also interested in diabetic-related medication use and variation in regional areas on the risk of dementia and therefore, I only took into consideration diabetic patients and the risk of incident dementia because my aim was to understand what are the factors within diabetic population which are risk factors for dementia diagnosis, because diabetes was the most prevalent comorbidity in my literature review of dementia studies and in my exploratory analysis of primary care data . Therefore, the limitation of this study is that I only used HES data for diagnoses and there was no matching of cohorts done based on age, sex, or diagnosis date. The study also lacks important clinical data such as smoking, alcohol consumption, BMI, genetic information which could be important factors affecting dementia risk.

4.9 Conclusion

This was an explanatory study to understand diabetic patient's risk of dementia. Much more research is needed to explore the underlying risk factors using linked EHRs data. The questions raised and the risk factors identified by this study have important implications and warrant further investigation.

Chapter 5: Multi-state modelling

In the traditional survival analysis, the model involves two states, alive and dead with only one possible transition between them. In chapter 4, I used a traditional Cox model because the outcome of interest was dementia diagnosis in the diabetic patients and the different factors affecting the hazard of this outcome.

In other situations, we may be interested in two or more intermediate events (transient) states, each of which represent a particular state of an illness or disease pathway. Therefore, we might be interested in factors affecting the transition rates between the different states ,i.e., the initial state, absorbing state(death) and all relevant intermediate states between. Multi-state models are used when at each time point the individuals occupy a particular state of an illness or disease pathway or when there are recurrent events for example multiple hospitalisations or episodes of infections. The traditional Cox model focuses on only one transition between the initial and the absorbing state and we cannot say anything about the entirety of an individual especially their probability in a specific state at certain point in time after the initial state. Multi-state models take this holistic approach and extend the Cox model to what happens after the first event, which then allows us to model the individual's transitions between different states [159]. These models allow us to model how a patient moves between states such hospital, discharge out, and institutionalisation. In multi-state model an individual must be in a specific state at a point in time and after the first transition the model consider the new hazard rate the individual is at risk. The following diagram illustrates multi-state models.

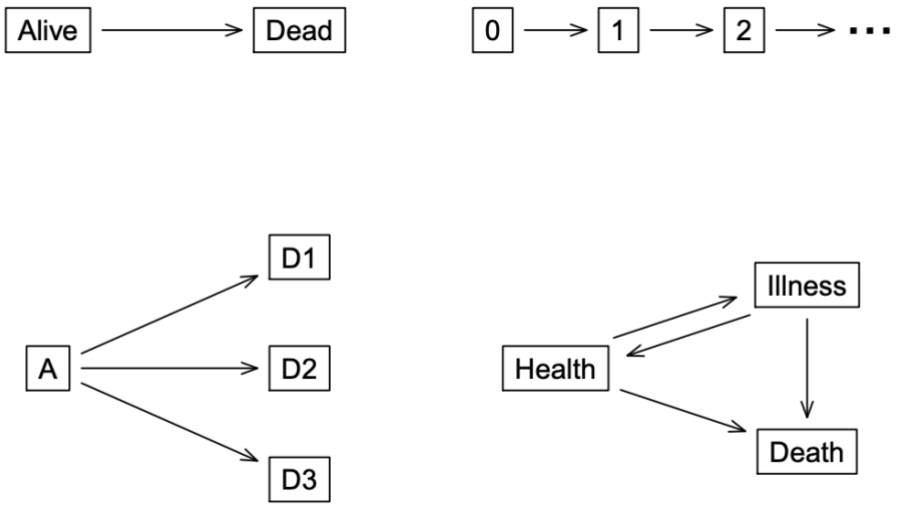


Figure 16 : Multi state models [160]

The upper left model is a simple survival model, the upper right is sequential events model, the lower left is competing risks where A is an initial state and D1, D2 and D3 are different competing events, and the lower right is a multi-state illness-death model.

A multi-state process is a stochastic process (X(t), t ∈ T) with a finite state space S = {1…..N}. Here T = {0,1,2, …}, if the process is in discrete time and T = (0, ∝) , if the process is in continuous time. X is the state of the process at time t and takes values from a finite state space S= {1….S). We can define the history of the process until time S, to be

$$H_s = \{Y(u); 0 \leq u \leq s\}$$

The history up until time s (Hs) will be generated with the evolution of the process over time and will consist of the observation of the process over the interval [0,t), the states visited over time and its transitions.

The multi-state process is fully characterised through transition probabilities or through transition intensities. The transition probabilities between states h and j can be calculated as follows:

$$P_{hj}(s,t) = P(X(t) = j \mid X(s) = h, H_{s^-})$$

for h, j ∈ S,

This is the probability of the process being in state "j" at time t, given that it was in state "h" at time s and conditional on the past trajectory before time s.

The transition intensities can be defined as

$$\alpha_{hj}(t) = \lim_{\Delta \to 0} \left(\frac{P_{hj}(t, t+\Delta t)}{\Delta t} \right)$$

Which represents the instantaneous hazard of moving to state j conditional on occupying state h.

5.1 Model Assumptions

Different assumptions can be made about the dependence of the transition rates (1) on time. These include.

5.1.1 Time homogenous models:

The intensities are constant over time, i.e., independent of t. Each transition probability, P_{hj} (s,t), in a time-homogenous model depends only on t—s, that is, P_{hj}(s,t) = P_{hj} (0, t – s).

5.1.2 Markov models

The transition intensities only depend on the history of the process through the current state. This simply means that the probability that a life now aged x will be in a particular state at any future time x+t depends only on the age x and the state currently occupied.

The behaviour of subsequent cycles or states depends only on its description in that cycle, i.e., the process has no memory for earlier cycles. For a multi-state model this means that, given the present state and the event history of a patient, the next state to be visited and the time at which this will occur will only depend on the present state. Another example to think of is that the future development of a 'diseased' individual at time t depends on the past only through the time elapsed since time origin (i.e., t and the fact that the individual is currently diseased) and not on the time span the individual has already been ill.

The Markov property assumes that the future depends on the history only through the present. Put another way, the Markov model adopts the homogeneity assumption by disregarding the pathway by which the previous event was reached. It is therefore time-homogenous (or time-stationary). In a multi-state model, it means that, given the present state and the event history of a subject, the next state that the subject will visit and the time at which that will occur will be dependent only on the present state that the subject is in.

Strictly speaking, only the clock-forward models can be Markov models.

For clock-reset models the Markov property cannot hold since the time itself depends on the history through the time since the current state was reached.

5.1.3 Semi-Markov models:

In Semi-Markov models, the duration times are dependent on the history of the process up to the present state and the duration or time since entry into that state. So, the future evolution not only depends on the current state h, but also on the entry time t_h into state h. This results in a sequence of embedded Markov models, often referred to as Markov renewal models or semi-Markov model. Therefore, we can consider intensity functions of the general form α_{hj} (t, t-th) or the special homogenous case α_{hj} (t-th). There are also extended Markov models, which involve for example, consideration of how the order of states visited influences transition rates, and there are models that allows the survival times to depend on the times at which earlier states have been entered. For example, in the model of transition rate from local recurrence and distant metastasis to death can vary according to a patient experiencing local recurrence before or after metastasis [159].

5.2 Stratified baseline hazard

A multi-state model is different from a standard Cox model because in multi-state models the baseline hazard is stratified for each of the allowed transitions within the model. In the traditional Cox model, since there is a single transition, only one baseline hazard $\alpha_0(t)$ is estimated. However, in a multi-state model, a separate $\alpha_{q0}(t)$ or baseline hazard is estimated for each transition q, where t is the time, an individual has been at risk [159].

5.3 Transition-specific covariates

In multi-state models there is a transition specific covariate effect, Zq. Here the q represents every possible transition in the model. The transition-specific covariate effect allows us to have a different effect by each covariate for each transition. For example, severe frailty may reduce the rate of institutionalisation but increase the risk of hospitalisation or death.

The hazard rate or intensity in a multi-state model is given by

$$\alpha_q(t) = \alpha_{q0}(t)e^{\beta^T Z_q} \quad (3)$$

$\alpha_{q0}(t)$ is separate baseline hazard in multi-state model and is estimated for each possible transition q, and t continues to refer to the amount of time that a subject has been at risk. And Zq is the transition-specific covariate effect where q is every possible transition in the model. We can also write the above equation as

$$\lambda_{ijk}\big(t \mid X_i(t)\big) = \lambda_{0,jk}(t)\exp\{X_i(t)^T\beta_{jk}\}$$

The right side of the above equation is the relative rate of regression and provides the relative rate (RR).

The cumulative hazard is also important and is estimated for transition i → j as:

$$H_{ij}(t) = \int_0^t q_{ij}\,(s)ds$$

Here S is the number of states within the model and u denotes all the transitions within some time interval (s,t] [161]. This gives us the state occupation probabilities, ,i.e., the probability of being in a state at time t. In the standard two state Cox-model we can compute the cumulative hazard as 1 minus the Kaplan-Meier estimate of survival probability. In multi-state models with competing events and several transitions we can use the Aalen-Johansen estimator to estimate the cumulative hazard which then gives us the probability of being in a state at a specific time, which is an important quantity of interest in multi-state models. The Aalen-Johansen estimator is a generalisation of the Kaplan-Meier estimator to multi-state models [160]. Aalen-Johansen is a product integral approach, and it gives us a vector p(t), which contains the probability of being in a state at time t. So, for each transition or event, the rate of transition is $l_{ij}(t)$,i.e., the fraction of individuals transitioning from state i to state j at time t, among those who are in state i just prior to time t. These transition probabilities are combined in a transition probability matrix P(s,t). We can estimate the transition probability matrix P(s,t) as

$$P(s,t) = \prod_{u \in (s,t]} \left(I + \Delta A(u)\right)$$

In the above (s,t] denotes the time interval and u indicates all event times in (s,t] and therefore, the matrix contains elements of the probability of transition from each state to every other state within the time interval (s,t][162]. The transition probabilities are based on the hazard rates, and these vary over time, these probabilities as well and therefore, the probability of an event at time t might be quite different than at time t+10.

5.4 Usefulness of multi-state models

Multi-state models are useful because they extend the traditional Cox models because these models can handle recurrent events and also offer the flexibility of predicting probabilities in some future time from the entry state. Multi-state models are of interest when we want to estimate the rate of transitions between states, i.e., how quickly will a patient move from one state to another, the probability of transition to a state, and the effects of different covariates on each of the possible transitions in a model. We will look at all of these factors which make multi-state models flexible and innovative.

5.4.1 Stratified baseline hazard

In multi-state models, the underlying baseline hazard is stratified for each of the possible transitions in the model. A separate baseline hazard, $\alpha_{q0}(t)$ is estimated for each transition q, where t is the time that the individual has been at risk. The traditional Cox model can take only one baseline hazard, $\alpha_0(t)$ because there is only one transition and therefore no q (transition) subscript. This allows us to accommodate the fact that the transition probability say from state 1 to state 2 is different than from state 1 to state 3. Stratification also is important when the model has repeated transitions for example, multiple hospitalisations. The rate or intensity at which a patient experiences the first hospitalisation may differ from the rate at which the individual experiences the 4th or 5th hospitalisation [163].

5.4.2 Transition-specific covariate estimates.

The multi-state model includes transition-specific covariates, Zq, where q indexes every possible transition in the model as shown in equation (3) above. This allows each covariate in the model to have different effect on different transitions in the model. For example, some x might decrease the risk of transitioning from hospital to a long-term institution

but might increase the rate of transitioning from hospital to the death state. So, as discussed above, the baseline hazard varies in multi-state models, but the effect of each independent covariate also does: the same covariate may exert a different effect on the rate of one transition from another. For example, increasing frailty may have a different effect on transitioning from hospital to a nursing home then discharge to patient's own home.

5.4.3 Transition probabilities.

One of the main advantages of multi-state models is their ability to accommodate predictions at different time points. We have several different states of interest in a multi-state model which we consider simultaneously, and this allows us to predict the state occupancy probabilities at a given time point. It also allows us to make dynamic predictions of the probability in a specific state as time elapses. Therefore, we can compare the prediction for the rate of transition to a state, e.g., death from the initial state or from an intermediate state in the model. In multi-state models, the estimation of probabilities of being in different states at various time points is important for understanding the dynamic transitions between states over time. These probabilities provide valuable information about the likelihood of an individual being in a certain state at a given time, which can be useful for clinical decision-making, prognosis, and health policy planning.

The estimated probabilities in multi-state models are typically obtained from non-parametric methods, such as the Aalen-Johansen estimator [164] , which is used to estimate the transition probabilities between different states at different time points.

The Aalen-Johansen estimator is a widely used non-parametric estimator in multi-state models. It estimates the cumulative transition probabilities between different states by considering the observed transitions, the number of individuals at risk of transitioning at each time point, and the duration of time intervals between observations. The estimator is based on counting the number of observed transitions and the number of individuals still at risk of transitioning at each time point, and it uses a stepwise approach to estimate the cumulative transition probabilities.

These estimated probabilities of being in different states at different time points provide important information for understanding the patterns of transitions between states in multi-state models and can be used to estimate various quantities of interest, such as survival probabilities, sojourn times, and other relevant clinical or policy measures.

In traditional time-to-event analysis we have only one transition but in multi-state models since we have more than two states and transitions, we need to aggregate the risk of each transition in the model. For example, in an illness-death model, individuals can transition to the death state from the "well" state or from the "illness" state and the multi-state model allow us to estimate an individual's transition probability for the death state taking into consideration the two transitions through which an individual can arrive in the death state. Therefore, by using multi-state model we can estimate an individual's probability of being in a specific state at a specific time t, t+1, t+2 and so on [165]. The transition probabilities help us estimate an aggregate transition probability in state when there are multiple transitions coming into a state which could be direct transition or indirect transition through another state. The transition probabilities are based on the following information.

(1) The fact that the baseline hazard is allowed to vary for each transition in the multi-state model, therefore, the current state occupied by an individual may have an

impact on their probability to transition to a subsequent state. For example, in a model where the states are hospital, discharge home, discharge to a nursing home and death, if the patients are being discharged quickly to their home but the transition from a nursing home to their own home is protracted, then an individual's probability of occupying the state "discharged home" will be different depending on the individual's current state.

(2) The second information needed to estimate the transition probabilities is the time frame for which we need the estimation. Do we need the transition probability to start from the initial time at the start of the study or after some time has elapse. For example, are we interested in the probability at age 60 or at dementia diagnosis or at age 80 or 10 years after the diagnosis of dementia.

(3) Covariates values of interest that is by providing a set of covariate values at which to obtain a prediction i.e., giving specific values of patient's covariates such as age, sex, and values for other prognostic factors to obtain the prediction for that specific profile of person. This allows us to predict the effect of a particular covariate in a specific transition but on the process as a whole.

In this chapter I described the theory of multi-state model, its assumptions and usefulness. I described how it is the right statistical model to use for our data since we have recurrent events, and we are interested in not only the hazard ratios but also in the probability of transition into a particular state over time. It also gives us the option to then predict the probabilities of transitioning into a particular state over time for specific profiles of patients. In the next chapter I will utilise the multi-state model for our dementia patients' data to estimate their rates of transition into the different states of our model and their probabilities over time.

Chapter 6: Multi-state model with Electronic Health Record data on dementia patients

In the previous chapter I described some of the aspects and benefits of multi-state modelling. In this chapter I will use the multi-state model to understand the influence of a dementia patient's risk profiles on hospital readmission, discharge to a long-term care institution (LTC) directly from hospital and death. I will be using the dementia patients identified in the primary care (CPRD) and using their hospitalisation record from the hospital episode statistics (HES) data. The influence of different patient's characteristics on these outcomes is largely unknown and by using a multi-state model, I will be able to jointly evaluate the impact of the different risk factors on the rate of transitions between hospital, LTC, and death.

Multi-state models are useful when analysing data on individuals who can move between different states over time. In the case of dementia patients, they can transition between different health states, such as being at home, being hospitalized, being institutionalized, and experiencing mortality.

Using a multi-state model to analyse these transitions allows for the simultaneous modelling of multiple outcomes, such as hospitalization, institutionalization, and mortality, in a single model. This has several advantages over analysing these outcomes separately:

Accounting for competing risks. When analysing multiple outcomes separately, it is important to account for the fact that individuals may experience different outcomes that are mutually exclusive. For example, if we analyse hospitalization and mortality separately, we need to account for the fact that an individual who is hospitalized may also die, and that these outcomes compete with each other. Multi-state models can account for this

competition between outcomes and provide more accurate estimates of the probability of each outcome.

Modelling the natural history of disease Multi-state models allow us to model the natural history of a disease, including the different stages of the disease and the transitions between these stages. In the case of dementia patients, we can model the progression of the disease and the different health states that patients may experience over time.

Estimating the effect of interventions Multi-state models can be used to estimate the effect of interventions on different health outcomes. For example, we can use a multi-state model to estimate the effect of a particular intervention such as home care on the probability of hospitalization, institutionalization, and mortality.

Overall, using a multi-state model to model dementia patients' hospitalizations, institutionalization, and mortality simultaneously allows for a more comprehensive analysis of the natural history of the disease and the effect of different interventions on multiple health outcomes.

Multi-state models have been used in heart failure studies using administrative data [166–168] however, in dementia studies, the use of multi-state models are scarce.

One study by Commenges et al (2004) [169], used a multi-state model for dementia, institutionalization and death using data from cohort studies. Their model starts with non-dementia patients living at home as the initial state and patients can then transition to dementia, institutionalised and death. They used simulated data to demonstrate the feasibility of using multi-state models and intended to use this model with the Paquid study data to study the relationship between dementia and institutionalization. The second study used home-dwelling persons with dementia from the Norwegian Registry of Persons Assessed for Cognitive Symptoms (NorCog) [170]. The objective was to estimate the transition times from dementia diagnosis to nursing home admission or death and to

examine whether sex, education, marital status, cognitive impairment, and dementia aetiology are associated with transition times.

Another study using a multi-state model to study dementia is a cross-sectional study on participants from the Mayo clinic study of aging without dementia at baseline to estimate four different age-specific biomarker state transition rates [171]. The aim was to estimate the rates of transition from a less to a more abnormal biomarker state by age in individuals without dementia and to assess the rates of transition to dementia from an abnormal state. The found that brain ageing is an inevitable acceleration towards worst biomarkers and clinical states.

The fourth study using a multi-state model used the same data as in the second study described above and endpoints were time to dementia and death [172] and using covariates such as amyloid burden, APOE-4, sex, education and cardiovascular/metabolic conditions (CMC). They estimated rates of incident dementia and mortality and found that high amyloid, APOE-4 and being a woman, old age and having two additional CMC were associated with the increased hazard of dementia. Men with or without dementia had a higher mortality than women.

All the above studies used data from cohort studies and were following non-dementia patients and using biomarkers data with single transitions between the different states of the models. Due to the nature of our data which is hospital admissions and discharge when a patient can be admitted several times to hospital and discharged to a care institution and admitted back to hospital several times, we need a model which can accommodate the multiple transitions between the states which can simultaneously model hospital admission, discharge out, discharge to a care institution and mortality. The model will allow us to estimate multiple readmission rates to both hospital and nursing home or

other care institutions directly from hospital. All the above mentioned studies did not use hospital admission and discharge data and did not consider multiple times a patient can transition between hospital and care institution which characterise dementia [173,174].

6.1 Introduction

Dementia is a progressive disorder which leads to loss of independence and eventually death. We have witnessed an increase in dementia cases in most developed countries due to better healthcare and decreased mortality. Due to the increased life expectancy more people are in a state of dependency and therefore, an increasing demand on formal and informal care. Formal care can be in an institutional setting or in the community and referred to as institutional care and community care respectively. In this study I used the term "institutionalisation" which refers to receiving care in a long-term residential facility. In England it includes NHS-run and privately run residential nursing homes. Similarly, care provided in the patient's own home is known as home care. Home care allows older people to stay at their own homes and keep their social network and independence which is good for their quality of life. Also, studies have reported that after accounting for the hidden informal care providers cost, that home care cost is significantly lower than care provided in the institutional settings[175,176]. This could be of interest to service providers and policy makers to save resources and make the care of dementia patients more efficient by increasing their quality of life by delaying institutionalisation and support them in their home for as long as possible.

Numerous factors are associated with dementia progression leading to institutional care admission and understanding of the factors affecting time to nursing home admission and related factors are of key importance to dementia patients and their families. In this analysis due to the nature of data at hand, I focus on discharge to institutional care from hospital. Discharge to a long-term care institution directly from hospital is something the

UK health policy advises against [177] but occurs quite often for example, according to the information service department of Scotland report from 2009-2015, 45% of people admitted to long-term care institutions came directly from hospital however, this number is down to 39% in their report in 2019[177]. Looking into the factors and predictors of long-term care discharge directly from hospital can potentially help service planning, and better understanding of the institutionalisation of dementia patients.

Different predictors of institutionalisation have been reported in the literature. These include age, sex, socioeconomic status, care giver characteristics, number of prescriptions, length of hospital stay, functional impairment, level of education [178–181].

We know that the diagnosis of dementia plays a vital role in institutionalisation, and it is well-documented[182] but the factors which mediate the rate of institutionalisation and death are not scrutinized. Our hypothesis is that within dementia patients the risk of institutionalisation and death is partially mediated by long hospital stay and admission due to injury. Also, we are interested to see if interventions such as home care, i.e., having an informal or formal carers at their usual place of living helps to delay institutionalisation and death in dementia patients. Identifying predictors of institution admission directly from hospital can potentially help planning, identification of targets to prevent long-term care admission and providing support to those at risk of this transition. It will also be of interest to service providers and policy makers.

6.2 Data and Methods

Dementia patients were identified in the primary care CPRD (Clinical Practice Research Datalink) and then their hospitalisation records were accessed in the hospital episode statistics (HES) data source. Patients diagnosed with dementia on or after 01/01/1998 were included if they had hospital records after their dementia diagnosis and had stayed in the hospital for at least a day. This gave a final 40,017 dementia patients to include in the analysis. The cohort selection is described in figure 17. Dementia patients were identified in CPRD using Read codes and in HES using ICD-10 codes listed in the appendix A tables 12 and 13. The information on institutionalisation from hospitals is derived from the HES data discharge destination codes as described in table 8 below. From the discharge destination codes we are able to tell whether a patient was discharged directly to a care institution such as a nursing home or other long-term accommodation where care is provided.

Table 9: Discharge destination codes used for determining institutionalization.

Discharge destination (disdest)	Code used in HES
NHS run nursing home, residential accommodation, or group home	54
Local authority residential accommodation where care is provided	65
Local authority foster care, residential accommodation where care is provided	66
Local authority home or care	69
Non-NHS run residential care homes (1996-1997)	85
Non-NHS run nursing home (1997 to 2007)	86
Non-NHS run hospice	88
Non-NHS run institution	89

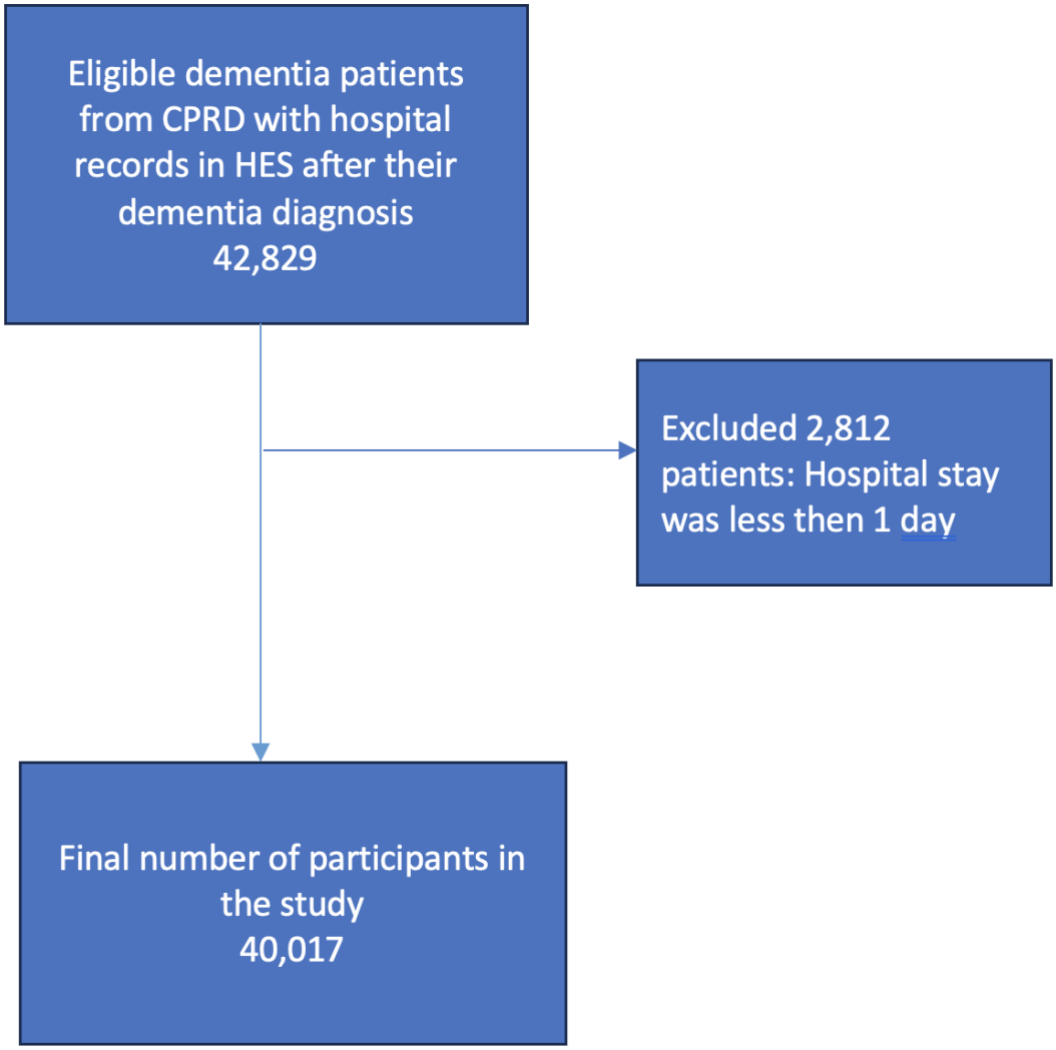


Figure 17 : Cohort selection process

6.3 Statistical methods

To characterise the association between patients' characteristics and hospital discharge to home or to a long-term institution and mortality, we adopt a multi-state Markov model describing how an individual moves between the different states. Suppose an individual is in state S(t) at time t and the next state to which he/she moves, and the time of change, are governed by a set of transition intensities $q_{rs}(t)$, r,s = 1,…., R. These intensities or hazards represent the instantaneous risk of transitioning from state r to state s. This could depend on the time t since the start of the process, patient characteristics z(t), and also could be the "history" of the process up to that point Ht which is the previous states visited by the individual and the time spent in them. Therefore,

$$q_{rs}(t) = \lim_{d\,t \to 0} P(S(t + d\,t) = s \mid S(t) = r)/d\,t$$

These are the elements of a R x R matrix Q(t) whose rows sum to zero, so that the diagonal entries are defined by $q_{rr}(t) = -\sum_{r \neq s} q_{rs}(t)$, and $q_{rs}(t)=0$ if a transition from state r to state s is not allowed and the patient will remain in state r as no transition is taking place.

When studying the transition intensities from hospital to a long-term care institution, death is a competing risk that cannot be ignored. People who die in hospital before being discharged to a long-term care institution will be censored in standard survival analysis. However, an important assumption of traditional time-to-events models such as Cox model is the non-informative censoring assumption, which means that censoring is independent of health status and assumes that the censored individuals are representative of those still at risk [183,184]. This assumption is violated if dementia patients are censored due to death. Therefore, to account for the competing risks and also the fact that we have

repeating transitions, I applied Markov multi-state model. Instead of the Kaplan-Meier, as in the Cox model, I will be using Aalen-Johansen estimator [164].

6.3.1 Structure of the model

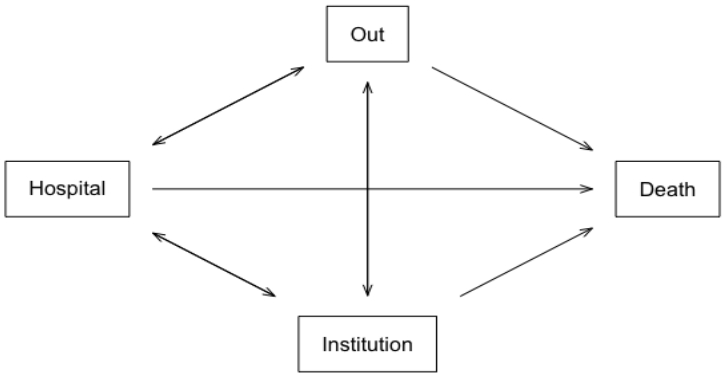


Figure 18: Model structure

The state space is made of all possible events described in the dataset.
Patients are in the "Out" state when they are either discharged out of hospital or out of institution before admitting to hospital again.
Institutionalisation in the above model refers to receiving care in a long-term care institution (LTCI)

There are 9 permitted transitions between the model states. Each patient starts in the hospital which is the index hospital admission after their dementia diagnosis date. From there they can be either discharged to an institution or discharged home or die while in hospital. They can be admitted to hospital again from an institution or once discharged out of hospital. Similarly, if they were discharged to an institution directly from hospital but their next hospital admission was not from an institution then that means they were discharged out of institution (institution → Out) and then admitted to hospital once out of institution. There transitions were determined using the HES discharged destination

and admission source codes. Each transition is a survival model, and the state the person is in, impacts the probability of the next transition (Figure 19). We can investigate the effect of different covariates on each transition in this multi-state model and so the model seeks to detect the impact of patient characteristics on the risk of moving within these states. Death is an absorbing state in the model, and it is a competing event with respect to all other transitions in the model.I used the survival package 3.3-1 [185] in R[186] and the data was reshaped to a counting process dataset using the tmerge function in the package. In the counting process dataset, there is one row for each transition for which the patient is at risk. All the patients in the model start as not living in institutional care after their dementia diagnosis (time 0) and the model then estimates the probability for each transition. Thus, unlike traditional survival models, which only estimate hazard ratios on a relative scale, a multi-state model also allows us to estimate the absolute probabilities.

The covariates used in the model include patient demographics data, prior use of anti-cardiovascular and anti-diabetic drugs because this will tell us if they are suffering from any cardiovascular diseases or are diabetic which could affect the hazard rate of different transitions in the model. Apart from that I used electronic frailty index (eFI) score, dementia subtype information, socioeconomic deprivations (Index of Multiple Derivation (IMD)) and also the hypothesized mediators which are admission due to injury and long hospital stay [187] (>= 12 days as per Alzheimer's society UK). The eFI score was estimated based on the methods developed by Clegg at al [188]. The methods use 36 deficits using diagnostic codes and then the frailty scores were used to classify patients into different frailty categories such as fit (eFI score of 0-0.12), mild frailty (eFI: >0.12-0.24), moderate frailty (eFI: >0.24 -0.36) and severe frailty (eFI: >0.36). For the socioeconomic deprivation, I used the index of multiple deprivation (IMD) 2015, which was grouped

into quintiles to group patients into IMD categories of 1 (least deprived), 2, 3, and 4 (most deprived).

Other covariates included in the model were: whether the patient was admitted back to hospital within 30 days or not, whether the admission was due to an injury and whether an individual was receiving home care or not. The last two covariates I hypothesise are mediators which influence the institutionalisation and mortality within dementia patients. Also, my aim was to look at the role of receiving home care in preventing/delaying institutionalisation and delaying mortality. The 30-day readmission, longer stay hospital admission, and admission due to an injury were all treated as binary variables, and they were also kept as time-varying variables as people could be admitted to hospital on multiple occasions for different reasons and for varying length of stay.

There are different packages in the available software's such as R, Stata and SAS, which can be used for multi-state modelling. Below is a comparison of packages available in R, Stata, and SAS for implementing multi-state models, along with their features and limitations. I used the Survival package in R because of its simplicity and the availability of function for data preparation, time-varying covariates adjustment and methods for competing risk availability.

Table 10: : A general overview of packages in different software tools for implementing multi-state models

Package	Language	Functionality	Limitations
R			
mstate	R	Fit multi-state models, visualize transitions	Limited support for time-varying covariates
msm	R	Fit multi-state models, simulate data	Limited support for competing risks, time-varying covariates
survival	**R**	Fit multi-state models, Cox models, Kaplan-Meier estimator	Limited flexibility in modeling complex multi-state scenarios, may require additional packages for certain functionalities, good for competing risks, time-varying covariates
JM	R	Joint modeling of longitudinal and time-to-event data	Requires additional packages for certain functionalities
flexsurv	R	Flexible parametric survival models	May require additional coding for complex multi-state models
Stata			
stjm	Stata	Joint modeling of longitudinal and time-to-event data	Limited flexibility in modeling complex multi-state scenarios
stcompeting	Stata	Competing risks analysis	May require additional coding for complex multi-state models
stcrreg	Stata	Cox regression for competing risks	Limited support for complex multi-state models
stpm2	Stata	Parametric multi-state models	Limited support for non-parametric models, time-varying covariates
SAS			
PROC PHREG	SAS	Cox proportional hazards regression	Limited support for multi-state models
PROC GENMOD	SAS	Generalized linear models	Limited support for multi-state models
PROC LIFETEST	SAS	Kaplan-Meier estimator, log-rank test	Limited support for multi-state models
PROC ICPHREG	SAS	Interval-censored proportional hazards model	Limited support for multi-state models

6.4 Outcomes

The outcomes from the multi-state model were to look at the relationship of different covariates with institutionalisation directly from hospital, rehospitalization and mortality in dementia patients over the age of 60 in the UK electronic health record. Each dementia patient was followed from their dementia diagnosis until their last hospital discharge or date of death. Our focus was on the overall hospitalisation in real-world setting to understand the total morbidity burden of dementia patients. I looked at the effect of covariates in the model on the rate of transitions between the different states of the model. Also, I looked at the probabilities in states for specific points in time in terms of absolute probabilities of being in a specific state and the effect of different covariates on these probabilities. The multi-state model also gave us the opportunity to estimate predictions for different profiles of patients, the chance of being in a certain state at some time point after dementia diagnosis. The model provides a description of the dementia patients admission-discharge dynamics looking at which covariates affect certain transitions and how they affect the relative risk (Hazard Ratios, HRs) as well as the risk, i.e., the instantaneous probabilities (absolute risk) of transitioning from one state to another. Understanding the factors affecting the rates of hospitalisation, institutionalisation and death are important for newly diagnosed dementia patients, their carers, and family. Also, the government and care providers want to delay institutionalisation and support dementia patients in their homes for as long as possible by providing home care. This model also helps us in understanding the effect of home care on institutionalisation and mortality. This information is important for policymakers and service providers because providing home care is significantly lower than institutional care even after accounting for the hidden cost of informal care [175,176,189]. In the statistical models, I was interested in the factors affecting all-cause hospitalisation, institutionalisation, and all-cause mortality. To improve our understanding of the factors affecting these transitions, and the prognosis

and impact of the different factors affecting these transitions, we need a comprehensive model which must include both the fatal and non-fatal clinical events. Therefore, a model which considers all possible events and transitions and can assess their dependence on important clinical factors is required.

6.5 Results

6.5.1 Descriptive analysis

A total of 40,017 dementia patients with hospitalisation records after their dementia diagnosis were identified. The characteristics of the cohort are described in table 9. The cohort consists of pre-dominantly female dementia patients with 66.8%, while males were only 33.2%. The mean age at diagnosis for men was 80 while for women it was 82 years. Compared to female dementia patients, male patients had a higher prevalence of vascular dementia (7.4% in men vs 4.5% in women), exposure to anti-diabetics (9.6% vs 5.9%) and exposure to anti-cardiovascular drugs (61% vs 54.8%) in the year before dementia diagnosis. The electronic frailty Index (eFI) score showed that more women were in the fit category than men (19% vs 16.5%) and more men than women were in the severely frail category (37.7% vs 35%). More men died then women during the follow-up period(86.2% vs 83.7%) and more women died in care homes compared to men (50.1% vs 42.6%) and more men died in hospitals/non-psychiatric places then women (37.4% vs 28.7%).

Table 11: Descriptive statistics of the cohort demographics

Characteristic	**Overall**, N = 40,017[1]	**Female**, N = 26,732[1]	**Male**, N = 13,285[1]	**p-value**[2]
Age at dementia diagnosis	81 (7)	82 (7)	80 (7)	<0.001
Home care				0.006
	8,582 (21%)	5,839 (22%)	2,743 (21%)	
Follow-up (Months)	55 (39)	58 (40)	50 (37)	<0.001
Dementia subtypes				<0.001
Unspecified dementia type	19,881 (50%)	13,614 (51%)	6,267 (47%)	
Alzheimer's disease	17,415 (44%)	11,682 (44%)	5,733 (43%)	
Rare dementia	526 (1.3%)	220 (0.8%)	306 (2.3%)	
Vascular dementia	2,195 (5.5%)	1,216 (4.5%)	979 (7.4%)	
Prior use of anti-diabetics				<0.001
	2,860 (7.1%)	1,587 (5.9%)	1,273 (9.6%)	
Prior use of anti-cardiovascular drugs				<0.001
	22,761 (57%)	14,657 (55%)	8,104 (61%)	

Characteristic	**Overall**, N = 40,017[1]	**Female**, N = 26,732[1]	**Male**, N = 13,285[1]	**p-value**[2]
Electronic Frailty Index (eFI)				<0.001
Fit	7,296 (18%)	5,102 (19%)	2,194 (17%)	
Mild frailty	10,128 (25%)	6,827 (26%)	3,301 (25%)	
Moderate frailty	8,231 (21%)	5,452 (20%)	2,779 (21%)	
Severe frailty	14,362 (36%)	9,351 (35%)	5,011 (38%)	
Place of death (POD)				<0.001
Care home	19,045 (48%)	13,384 (50%)	5,661 (43%)	
Home/Elsewhere	8,126 (20%)	5,566 (21%)	2,560 (19%)	
Hospital/Non-Psychiatric	12,629 (32%)	7,667 (29%)	4,962 (37%)	
Psychiatric Places	217 (0.5%)	115 (0.4%)	102 (0.8%)	

[1]Mean (SD); n (%)

[2]Wilcoxon rank sum test; Pearson's Chi-squared test

Table 12: Descriptive analysis of all cause rehospitalizations

Characteristic	Overall, N = 40,017[1]	Index Hospitalisation, N = 14,496[1]	One Rehospitalisations, N = 10,128[1]	Two Rehospitalisation, N = 6,136[1]	Three or more Rehospitalisations, N = 9,257[1]	p-value[2]
Gender						<0.001
Female	26,732 (67%)	9,842 (68%)	6,866 (68%)	4,090 (67%)	5,934 (64%)	
Age at dementia diagnosis	81 (7)	82 (7)	81 (7)	81 (7)	80 (7)	<0.001
Home care	8,582 (21%)	2,550 (18%)	2,019 (20%)	1,416 (23%)	2,597 (28%)	<0.001
Follow-up (Months)	55 (39)	43 (36)	52 (37)	60 (38)	73 (41)	<0.001
Dementia subtypes						<0.001
Unspecified dementia type	19,881 (50%)	7,348 (51%)	5,057 (50%)	2,998 (49%)	4,478 (48%)	
Alzheimer's disease	17,415 (44%)	6,244 (43%)	4,413 (44%)	2,738 (45%)	4,020 (43%)	
Rare dementia	526 (1.3%)	188 (1.3%)	129 (1.3%)	69 (1.1%)	140 (1.5%)	
Vascular dementia	2,195 (5.5%)	716 (4.9%)	529 (5.2%)	331 (5.4%)	619 (6.7%)	
Prior use of anti-diabetics						<0.001
	2,860 (7.1%)	910 (6.3%)	710 (7.0%)	443 (7.2%)	797 (8.6%)	
Prior use of anti-cardiovascular drugs	22,761 (57%)	8,531 (59%)	5,697 (56%)	3,414 (56%)	5,119 (55%)	<0.001

Characteristic	**Overall**, N = 40,017[1]	**Index Hospitalisation**, N = 14,496[1]	**One Rehospitalisations**, N = 10,128[1]	**Two Rehospitalisation**, N = 6,136[1]	**Three or more Rehospitalisations**, N = 9,257[1]	**p-value**[2]
Electronic Frailty Index (eFI)						<0.001
Fit	7,296 (18%)	2,209 (15%)	1,809 (18%)	1,247 (20%)	2,031 (22%)	
Mild frailty	10,128 (25%)	3,528 (24%)	2,646 (26%)	1,602 (26%)	2,352 (25%)	
Moderate frailty	8,231 (21%)	3,220 (22%)	2,092 (21%)	1,226 (20%)	1,693 (18%)	
Severe frailty	14,362 (36%)	5,539 (38%)	3,581 (35%)	2,061 (34%)	3,181 (34%)	
Admission due to injury	15,422 (39%)	3,368 (23%)	3,892 (38%)	2,810 (46%)	5,352 (58%)	<0.001
Place of death (POD)						<0.001
Care home	19,045 (48%)	6,852 (47%)	4,873 (48%)	2,968 (48%)	4,352 (47%)	
Home/Elsewhere	8,126 (20%)	3,132 (22%)	1,998 (20%)	1,179 (19%)	1,817 (20%)	
Hospital/Non-Psychiatric	12,629 (32%)	4,411 (30%)	3,204 (32%)	1,961 (32%)	3,053 (33%)	
Psychiatric Places	217 (0.5%)	101 (0.7%)	53 (0.5%)	28 (0.5%)	35 (0.4%)	

[1]n (%); Mean (SD)

[2]Pearson's Chi-squared test; Kruskal-Wallis rank sum test

The clinical characteristics of the cohort according to the number of hospitalisations after the index hospital admission (First hospital admission after dementia diagnosis) are shown in Table 10. Overall, the readmission rate was associated with increasing frailty and admission due to injury, which increased the rehospitalisation rate from 23.2% to 57.8%. People with home care had more of readmissions, probably due to their decreased mortality.

6.5.2 Probability in state Ps(t) or state occupation probability (SOP)

6.5.2.1 SOP Institutionalisation state

For the survival curves or probability in state curves I use the Aalen-Johansen estimator to compute the state occupation probabilities, which describes the distribution of patients in a specific state at a specific time point during the follow-up. It, measures how many people are in each state at any given time. From a clinical perspective, SOP can be used to estimate the probability of a patient being in a particular state at a given time point, and how this probability changes over time. We can use SOP for example, in cancer progression, to estimate the probability of being in a particular stage of the disease at a given time and how this probability might change with different interventions. This can help clinicians to make decisions about the most appropriate treatment strategy for individual patients. In our example, we can estimate and plot the SOPs for the different states to check the difference in the proportion of people in a particular state at different age or time point for men and women and for the home-care intervention.

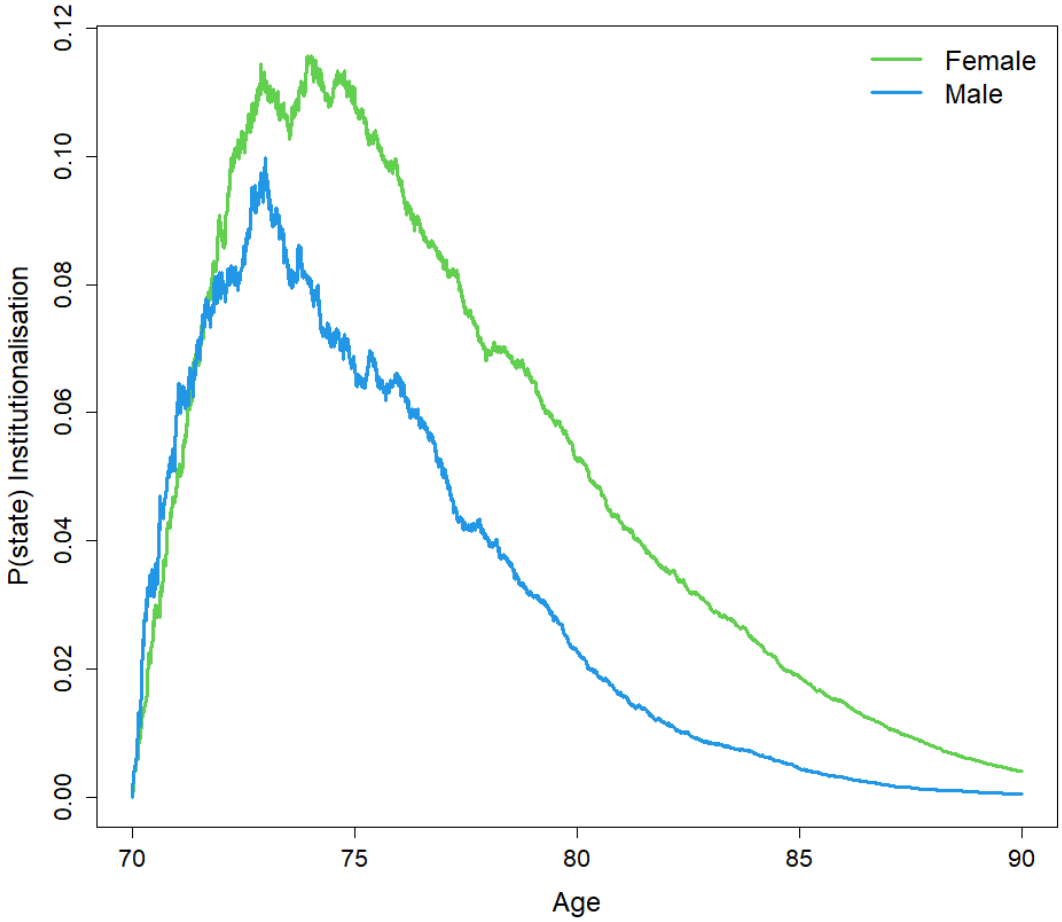


Figure 20: Institution State occupation probability (SOC) for men and women

Figure 20 above shows the probability of being discharged directly from hospital into an institution over time for men and women with dementia. Starting at age 70, we can see that men are discharged slightly quicker than women of similar age and with increasing age more women are discharged to institutional care then men and the probability of men being discharged to institutional care decreases quickly after age 75, due to their higher mortality.

Similarly, if we look at the probability of discharge directly from hospital to an institution for dementia patients with and without home care, the probability of being discharge to an institution is higher for those with no home care which is not surprising. The

probability of being discharged to an institution peaked at 8% and 12% for dementia patients with home care and without home care respectively. This probability remains high even after age 80 for people who have home care which shows their lower mortality rate compared to those without home care (figure 21). The curve goes up when someone is discharged from hospital to an institution and down when someone leaves this state (i.e. to transition back to hospital, home state or death state).

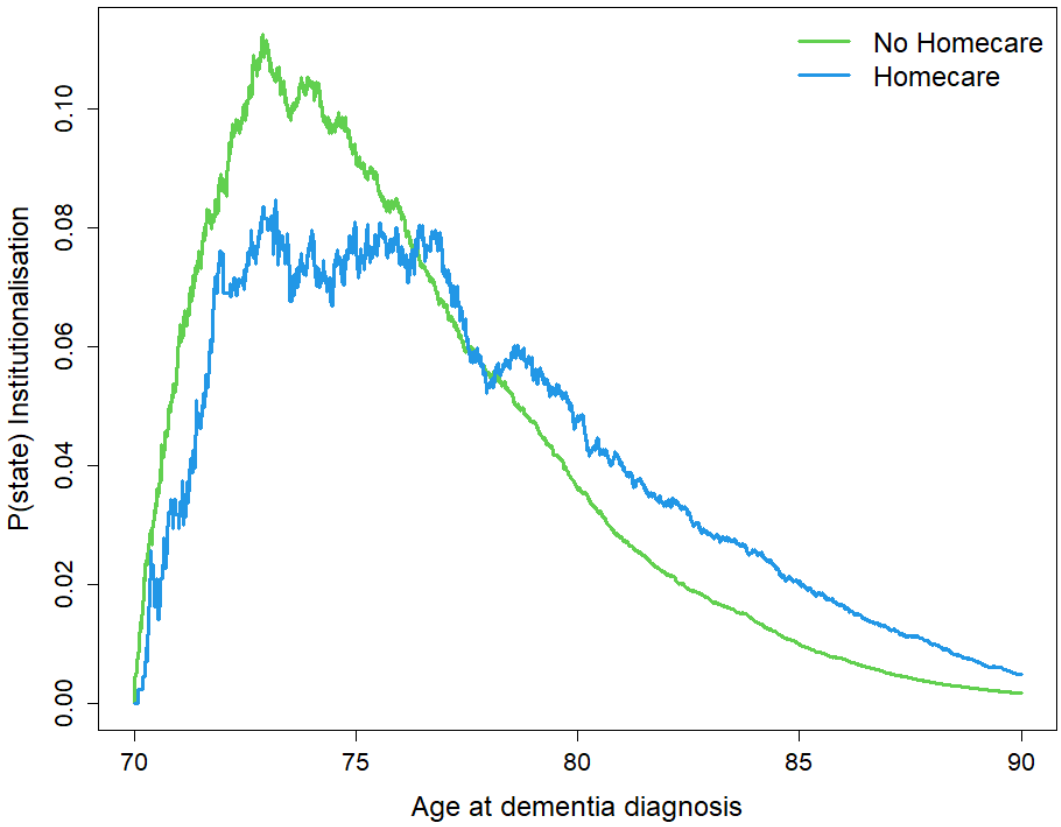


Figure 21 Institution State occupation probability (SOC) for dementia patients with and without home care

6.5.3 SOP for death state

The state occupation probabilities (SOPs) for the death state for male and female dementia patients at different ages is given in figure 22.

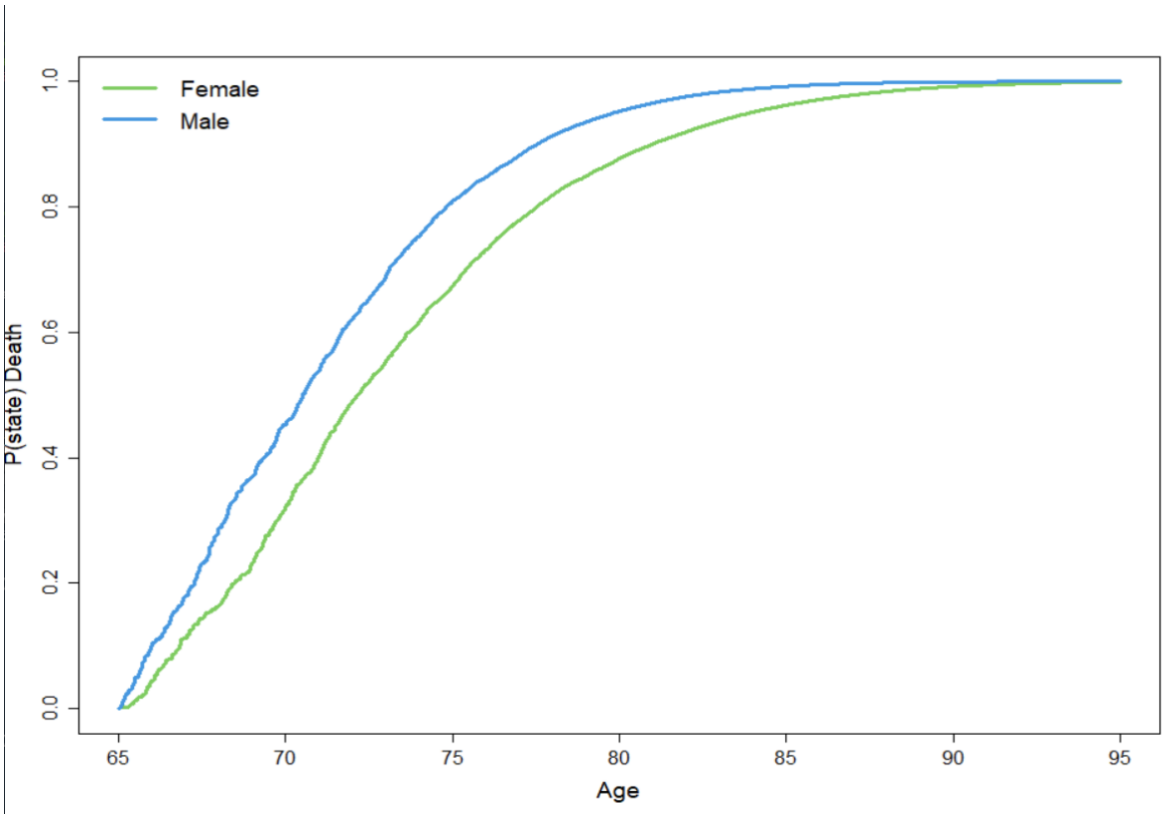


Figure 22: Probabilities in state Ps(t) in the death state for men and women

Probability of death increased with increasing age for both men and women however, men were more likely to be in the death state compared to women. At age 70 the probability of being in the death state for men was 45% compared to 32% for women. At age 80 the SOP of death state for men was 95% and for women it was 87.7% and at age 90 the probability is almost the same for both men and women.

6.5.4 Dementia patients with exposure to anti-cardiovascular medications

Dementia patients who suffer from cardiovascular conditions and take anti-cardiovascular drugs had a greater hazard rate of mortality (figure 23), and therefore they most likely to die within hospital and therefore, their probability of occupying or transitioning to other states is lower. The probability of dementia patients having exposure to anti-cardiovascular drugs peaked at 7% between 2-4 years after dementia diagnosis, while people without cardiovascular conditions were less likely to die and therefore the proportion of patients were lower in the death state but with increasing age or time after dementia diagnosis as they were living longer were discharged in greater proportion to the institution state.

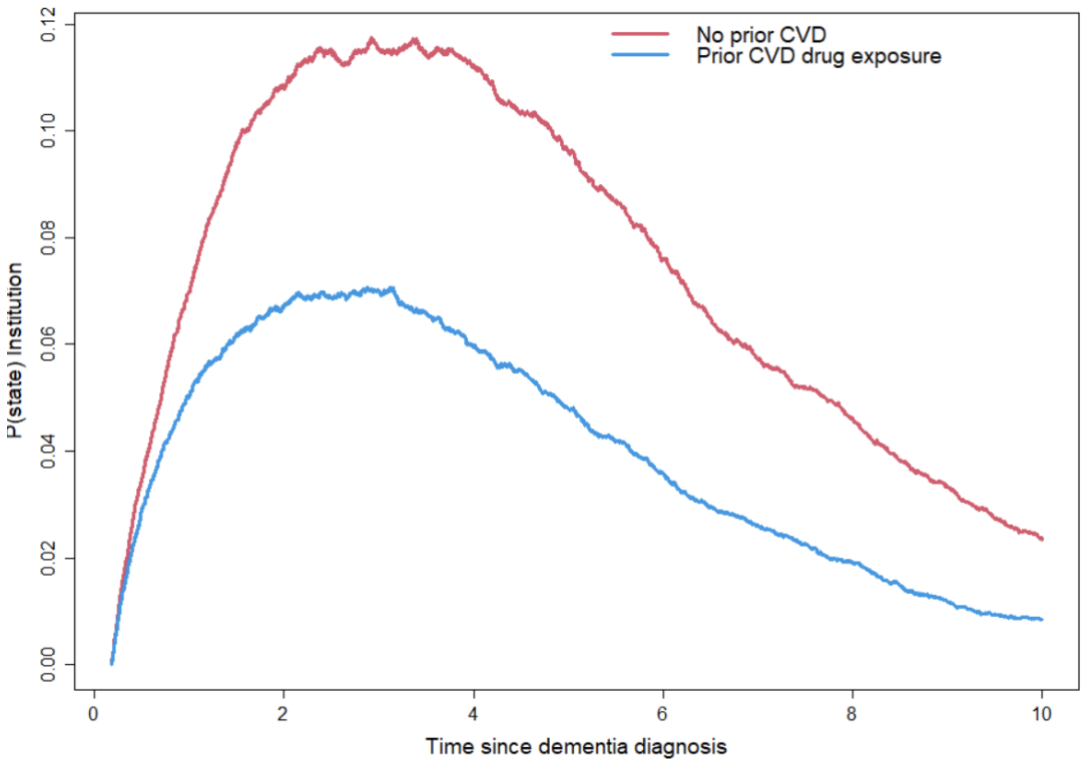


Figure 23: Institution state occupation probability for dementia patients with and without exposure to anti-cardiovascular medications.

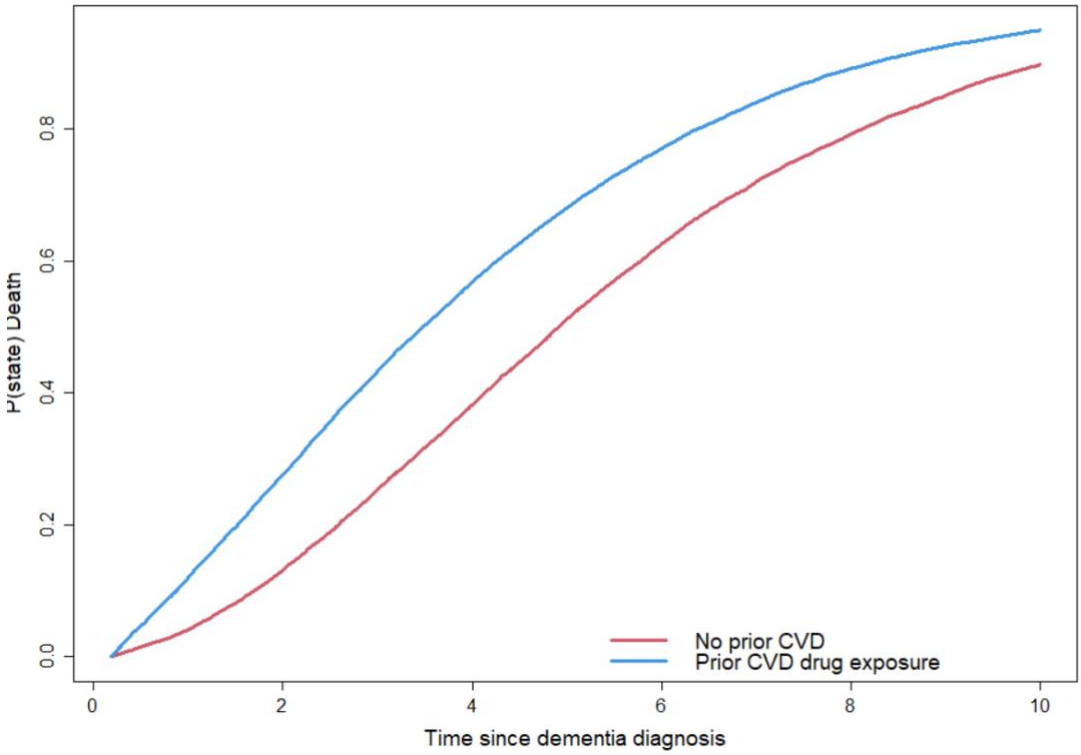


Figure24: Death state occupation probability for dementia patients with and without exposure to anti-cardiovascular medications.

However, those dementia patients who had cardiovascular complications but had home care as well had a similar probability of death as of those patients who did not have cardiovascular complications. We saw in the above figure (figure 24) that patients who had cardiovascular complications were taking anti-cardiovascular drugs had a greater probability of ending up in the death state, however, figure 25 shows that after age 75 years, having home care with cardiovascular complications reduces probability of death and give them the rate of those with no cardiovascular complications.

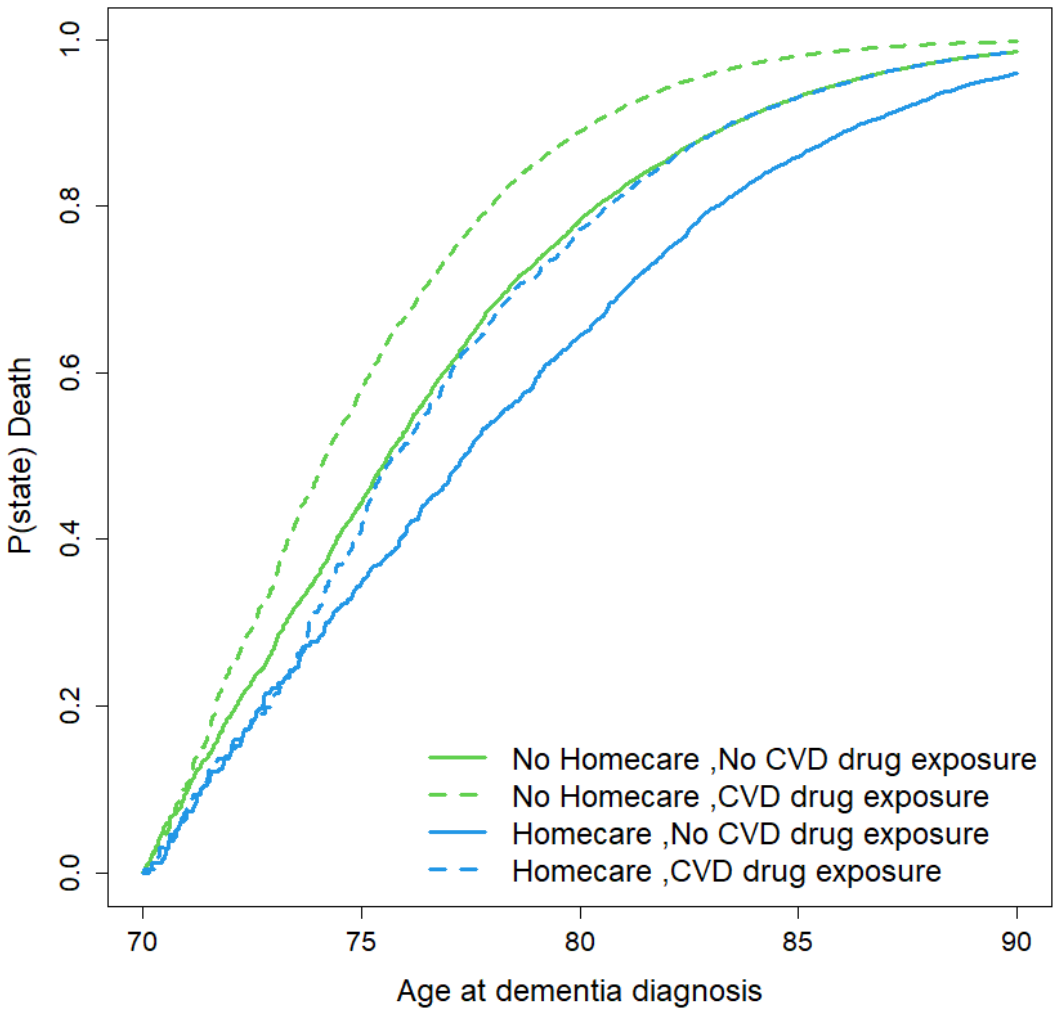


Figure25: Death state occupation probability for dementia patients with and without home care and cardiovascular complications

The above graphs show the probabilities in state or the state occupation probabilities (SOP) which are overall probabilities for the cohort without any assumptions about individual characteristics or covariates. The above graphs show how the probability of being in the institution state or in the death state changes with time for men and women and dementia patients with and without home care. These plots can be used to make clinical decisions such as looking at the prognosis or progression of dementia patients with and without home care and different age groups.

6.6 Multi-state Cox model

The multi-state model described in figure 18, measures the relative hazard or the transition intensity which measure how quickly one move from one state to another in instantaneous time. The multi-state Cox model focuses on the transitions or the arrows in the model and each arrow is a separate Cox model in itself. This provides us with the admission-discharge dynamics, pointing out which covariates affect certain transitions and what effect it has on the relative risk as well as on the risk or the instantaneous probability of transitioning from one state to another.

The multi-state model assumes the patients to be in one of the five states ,i.e., initial (S0) state, hospital, out, institution and death. The model allows us to detect the impact of patient characteristics on the intensity of transitioning between these states. The death state in the model is an absorbing state.

The hazard ratios (HR) and coefficient effects of all the covariates on each transition in the model are given in table 15 in appendix A. These HR estimates represent the instantaneous risks of transition from one state to another. I will describe the effect on the instantaneous hazard of transition by the different covariates in the model.

6.6.1 Home care

The relative magnitude of receiving home care on the hazard of being discharged to institution was higher directly from hospital but transition to death state and transition to institution once discharged from hospital was lower compared to those who were not receiving home care. In the literature the conceptualisation and measurement of home care is somehow imprecise, and most studies use living alone or the existence of living children as a proxy variable for receiving home care [190–192]. Arguably, home care is

associated with the availability of informal or formal care [193–195]. In this study I have used an indicator of whether the patient had a Read code suggesting the availability of informal or formal home care or Read codes in CPRD and ICD-10 code in HES suggesting living alone as a proxy for the availability of home care. Our model's estimates for the influence of home care on dementia patients' discharge to care institution directly from hospital are slightly higher (HR 1.07; CI 1.01 – 1.14) compared to dementia patients not receiving home care.

However, their transition to institutional care from outside of hospital was significantly lower than people with no home care (HR 0.76, CI 0.65-0.88) and their transition intensity to discharge out of hospital was also higher (HR 1.29, CI 1.23-1.37). This shows that people with home care were less likely to die inside hospital as their transition rate going out of hospital was higher and their transition to the death state directly from hospital was lower (HR 0.75, CI 0.70- 0.81) but also their transition rate to death state from their homes (Out to death state) and death inside institution (Institution to death state) were lower with HR 0.68 and CI 0.66-0.71 and CI 0.87 and CI of 0.81-0.93 respectively.

It is also unlikely that this could be due to the compounding effect that people who receive home care may be in good health or be less dependent as I have controlled for exposure to cardiovascular medication, frailty and age in the model. All this indicates that receiving home care does have an effect on mortality of dementia patients and they were less likely to die inside hospital or outside and their transition rate of discharge directly from hospital to their homes were higher which means they were spending less time inside hospital. It is also unlikely that this could be due to the compounding effect that people who receive home care may be in good health or be less dependent as I have controlled for comorbidity burden and frailty and age in the model.

Table 13: Effect of home care on dementia patients transitions between the model states.

Home Care	Estimate	P value	CI low	CI high
Subsequent hospital admission from Out (State 3 -- > State 2)	1.06	1.66E-04	1.03	1.09
Readmission from Institution (State 4 to 2)	1.03	6.32E-01	0.92	1.14
Discharge Home (2 to 3)	1.29	4.38E-20	1.23	1.37
Discharge to Institution(State 2 to 4)	1.07	3.29E-02	1.01	1.14
Home to Institution (state 3 to 4)	0.76	2.76E-04	0.65	0.88
Death inside hospital (state 2 to 5)	0.75	4.09E-15	0.70	0.81
Death outside hospital (state 3 to 5)	0.68	3.72E-82	0.66	0.71
Death inside institution (state 4 to 5)	0.87	2.78E-05	0.81	0.93

6.6.2 Age

With increasing age hospital readmission decreased, hospital discharge to patients' usual place of living decreased, however, the rate or intensity of discharge from hospital directly to care institution increased with increasing age. The readmission rate from patient's usual place of living and readmission to hospital once they were discharged to an institutional care for the age group 85+ for example was HR 0.96, CI 0.92-1.00 and HR 0.85, CI 0.74-0.97 respectively. This shows that older individuals were spending more time in institutional care and inside hospital and therefore, they were more likely to die inside hospital or a care institution. This was also evident from the higher rate of mortality within hospital (HR 2.16, CI 1.96-2.38) and within care institution (HR 1.96, CI 1.79-2.14) for dementia patients within the age group 85+ for example.

6.6.3 Socioeconomic deprivation

Using the English indices of deprivation (IMD) 2015 as a covariate we looked at the effect of socioeconomic deprivation on dementia patients' rate of transitions between the different states. The results show that with increasing socioeconomic deprivation, the rate of rehospitalisation also increased. The hazard rate for getting admitted to hospital again once discharged from outside for the IMD4 vs IMD1 was 1.17, CI 1.13-1.22 and getting admitted to hospital from a care institution 1.35, CI 1.20-1.53. Increasing socioeconomic deprivation was also associated with increasing rate of institutionalisation both directly from hospital and once discharged out of hospital . These hazard rates were: 1.30, CI1.20-1.40 for the IMD4 vs IMD1(least deprived) for the rate of institutional discharge directly from hospital and 1.71, CI 1.44-2.04 for the transition rate of going to institutional care once discharged home from

hospital. The hazard of death inside hospital significantly increased with increasing socioeconomic deprivation with an HR of 1.23, CI 1.12-1.34, however, the rate of death outside hospital and inside a care institution was greatly reduced. This hazard rate for the death inside care institution was 0.84, CI 0.77-0.90. This shows that dementia patients living in most deprived area were spending more time in hospital and therefore were more likely to die inside hospital and also their rate of hospital admission from a care institution was higher.

6.6.4 Dementia subtypes

Patients with a recorded dementia diagnosis as Alzheimer's dementia (AD) had a higher rate of hospital discharge both to their usual place of living and to a care institution. The transition intensity or rate for discharge out had an HR 1.06, CI 1.01-1.13 and for discharge to a care institution 1.12, CI 1.06-1.18 compared to the dementia patients identified as having unspecified dementia. Only AD and vascular dementia (VaD) patients had a higher rate of hospital readmission from a care institution and their HR, and 95% confidence intervals were 1.20, CI 1.10-1.31 for AD and 1.41, CI 1.19-1.68 for VaD for the transition of hospital admission once they were already in a care institution.

Dementia subtypes did not show any significant association with the rate of mortality.

6.6.5 Exposure to anti-cardiovascular and anti-diabetic drugs

In our systematic review of dementia progression studies and the exploratory analysis of our electronic health databases, cardiovascular diseases and diabetes were the top comorbidities within dementia patients. Also, it is reported in the published literature, and I also found some evidence of reduced dementia incidence in people who were taking anti-cardiovascular and anti-diabetic medications. Therefore, I included exposure to anti-cardiovascular and anti-diabetic medications in the year before dementia diagnosis to see if there is any effect of the

exposure to these medications on rehospitalizations, discharge to institutional care or mortality within dementia patients.

The rate of hospital readmission once discharged out to their usual place of living was slightly lower for dementia patients who were taking anti-cardiovascular drugs with HR 0.95, CI 0.92-0.98 and was higher for dementia patients with exposure to anti-diabetic drugs meaning dementia patients with diabetes with HR 1.11, CI 1.05-1.17. Dementia patients with cardiovascular conditions were more likely to die inside hospital HR 1.28, CI 1.19-1.38 and they also had a higher rate of death inside institutional care settings HR 1.22, CI 1.14-1.30. This shows the effect of cardiovascular conditions on dementia patients and increases their hospitalisations and mortality rate.

6.6.6 Frailty

Frailty has been associated with a range of adverse health outcomes, including increased hospitalizations, institutionalizations, and mortality in patients with dementia[196–198]. I estimated an electronic frailty score (eFI) using the method described by Clegg et al [188] where they used 36 deficits and a frailty index score was estimated. The eFI score was then categorised into fit (eFI 0-0.12), mildly frail (eFI >0.12 – 0.24), moderately frail (>0.24-0.36) and severely frail (>0.36). Increasing frailty within dementia patients increased the rate of rehospitalisation both from care institution and patients' homes and the model estimates were HR 2.31, CI 2.00- 2.66 and HR 1.80, CI 1.72-1.89 for patients with severe frailty for rehospitalisation from care institutions and patients usual place of living respectively. However, increasing frailty was also associated with increased rate of hospital discharge to both patients' own home and care institutions, which means they were spending less time in hospital. Descriptive analysis of the data revealed that patients with severe frailty were mainly aged 80 and above and they mostly had home care which could be the reason of them being

discharged to their homes and care institutions. The HR for hospital discharged directly to care institution was 1.65, CI 1.51-1.81 and to their usual place of living 2.05, CI 1.88-2.23 for dementia patients with severe frailty.

6.6.7 Thirty-day readmission

I use 30-day readmission as time-dependent covariate in the model to estimate its effect on discharge from hospital to institutional care, hospital readmission and death. It was used as a binary variable if a hospital admission was within 30 days from the previous hospital discharge. The effect of this variable was not statistically significant for institutional care transition from hospital or death inside hospital, however, their rate of institutionalisation once discharged from hospital alive and death outside hospital was higher. They were also at increased risk of hospital readmission both from the out state and from the institution state. These results are interesting because these patients had a 9% higher rate of being discharged out of hospital (HR 1.09, CI 1.03-1.15) which could be due to the hospitals effort to decrease the hospital length of stay, but this increases the readmission as we see the higher HR from both the Out (HR 1.62, CI 1.56-1.68) and institution state (HR 1.18, CI 1.07-1.31). Death inside hospital wasn't significant but death outside hospital once discharged home or to institutional care were significant and for the rate of death once discharged out was HR 1.32, CI 1.27-1.38 and within institutional care it was HR 1.09, CI 1.02-1.16.

These estimates shows that hospital readmission within 30 days is associated with their increasing discharge rate from hospital (shortening length of stay) which in turn increases their rate of institutionalisation and death once out of hospital and of course readmission to hospital.

6.6.8 Admission due to injury

The effect of hospital admission due to an injury was checked for hospital discharged to institutional care and death inside hospital. Dementia patients who had hospital admission due to an injury were more likely to be institutionalised after hospital admission compared to those whose admission was not due to an injury (HR 1.71, CI 1.45 -2.01). Patient with injury admissions also had a higher rate of death inside hospital (HR 1.30, CI 1.09- 1.56).

6.6.9 Long hospital stay

The effect of long hospital stay was checked for discharge to institution directly from and death inside hospital. The long stay variable was defined as a binary variable of whether the patient's stay in hospital >= 12 days was or not. The 12 days number is based on the Alzheimer's society UK report published in 2017/18 [2]. With long hospital stay the likelihood of discharge to institutional care directly from hospital with HR of 1.34 (CI 1.22-1.48). Patients who had long hospital stays were also at increased risk of death inside hospital death (HR 1.20 , CI 1.08-1.34).

6.7 Predictions from the Multi-state model

The multi-state model also gives us the opportunity to use the results from the multi-state Cox model to obtain predictions for the probabilities in a specific state at a certain time after dementia diagnosis for patients with a given set of covariates.

---

[2] According to the Alzheimer's society UK, the mean length of hospital stay for dementia patients in 2017/18 was 11.4 days

(https://www.alzheimers.org.uk/sites/default/files/2020-01/Hospital%20Admissions%202012-18%20for%20people%20with%20dementia%20Alzheimer%27s%20Society%202020.pdf)

The probabilities in state obtained from a multi-state Cox model can provide valuable information about the progression of a disease or the likelihood of experiencing a particular event over time. By incorporating covariate information, these models can provide personalized predictions for individual patients, allowing for more targeted and effective interventions [199–201].

Predictions from a multi-state Cox model provide an estimate of the probability of being in a specific state at a specific time, based on the patient's characteristics (covariates) and the estimated hazard ratios from the model. These predictions can be used to estimate the probability of transition to a certain state, given a set of covariates for an individual patient.

This type of prediction is important for several reasons. First, it allows clinicians and researchers to estimate the prognosis of individual patients, based on their unique characteristics. This can be useful for making treatment decisions and for counselling patients and their families about the potential outcomes of their disease.

Second, predictions from a multi-state Cox model can also be used for causal inference. For example, if a researcher is interested in estimating the effect of a certain treatment on the probability of transitioning from one state to another, they can use the model to estimate the hazard ratios and predict the probabilities of transition for patients with and without the treatment. This can help to estimate the causal effect of the treatment on the outcome of interest. In my model I used the availability of home care as an intervention to see its effect on the probability of transitioning to institutional care and death state in the model while controlling for other relevant prognostic factors. As an example, patient, I took a patient in the

age group 75-85, diagnosed with AD, who was also exposed to anti-diabetic and anti-cardiovascular medications and with moderate frailty, no admission due to injury and with an IMD (Index of Multiple Deprivation) of 4 and I estimated the probability curves for male and female dementia patients with and without home care. For these groups of patients, with the above characteristic profiles, the risk of transitioning into institution state and death was higher for men than women, especially for those who did not have home care. Women with home care had the lowest risk of going into hospital after dementia diagnosis especially in the first two years after diagnosis, although this difference is very small, and this could be due to the nature of the data as these are elderly dementia patients and have very similar hospitalisation rates. However, although these data do not give us the full picture of the impact of home care, it still shows that home care has an impact on these transitions. Similarly, the hospital discharge (Out state) shows that men and women with home care were discharged at a quicker rate and were spending less time in hospital compared to those with no home care. Hospital discharge directly to institutional care was also higher for dementia patients with no home care and was higher for women than men because of their lower mortality compared to men. Men had a lower rate of institutionalisation directly from hospital because of their higher mortality rate however, among men those with home care had a lower rate compared to men with no home care. This was true for women as well instead that women were taking an extra year, and their probability of institutionalisation was higher for those with no home care until 3 years after diagnosis and decreased after that due to higher mortality rate of women with no home care. This shows the impact of home care on both institutionalisation and death. Similarly, the probability of transition to the death state was higher for patients with no home care. It was higher for men compared to women however, men who had home care had the same rate as women with no home care which mean home care gives men the rate of women mortality rate which was lower.

These types of predictions can be checked for any profile of patient and allow us to estimate more personalised prediction which helps with effective management of these patients.

6.8 Discussion

By using a Multi-state Cox model, I showed the entire process of dementia patients' hospitalisation, discharge home, institutionalisation and death can be modelled simultaneously using one model. Exploring the effect of different covariates on each transition in the model allowed us to compute overall transition probabilities taking into consideration the repeated transitions and competing risks. Dementia is associated with institutionalisation, however, within dementia patients there are other factors affecting the rate of hospitalisation, and institutionalisation which has rarely been examined.

In my model I focused on institutionalization from hospital as it is the major source of admission to care institutions. The model also looked at discharge out of hospital, rehospitalizations and mortality and the effects of admission due to injury, long hospital stays, and other prognostic factors important in dementia prognosis. Admission due to injury and long hospital stays are common in dementia patients but I wanted to enhance our understanding of the effects of these factors within dementia patients which has not been explored before.

Due to data constraints, I am not able to say to what extent home care influences the risk of institutionalisation, rehospitalisation, and mortality but our results suggest that home care was associated with a lower rate of mortality, and hospital discharge to patient's homes, however, it was associated with a higher rate of hospital discharge to institutional care in terms of hazard ratios. The hazards ratios tell us about the relative risk of transition at a specific point in time from a particular state. However, the absolute risk or the probability of getting institutionalised between those with and without home care which is given by the state occupation probability (SOP) described in figure 21, shows that dementia

patients with home care were delaying their institutionalisation and because of their lower mortality they were spending more time later in their life in an institutional care setting. Adequate support provided at home is very important for older people's independence and quality of life and our results suggest an important role of home care among dementia patients.

Our data come from large administrative datasets of primary care data from GP clinics across UK (CPRD), and HES data which contains hospital admissions and discharge information and therefore have an advantage over using other data such as survey data because the prevalence of dementia in these datasets is ~ 28% of the 850,000 dementia patients in the UK [85]. Also, using administrative data have advantage over other data such as data from specific clinical settings and with limited variations. In contrast the data I used comes from over 2000 UK primary care practices containing patients with different socioeconomic backgrounds and comorbidity profiles. Multi-state model has not been previously used to study dementia patients' outcomes such as hospitalisation, institutionalisation, and mortality. For the first time I demonstrated the effect of mediators such as admission due to injury, long hospital stays, 30-day readmission within dementia patients on their risk of institutionalisation directly from hospital, and their rate of mortality using primary care and secondary care data within a single model. Other well-known predictors of institutionalisation such as old age, gender, socioeconomic status, frailty which are widely reported in the literature were also looked at however, we know little about an effective intervention that reduce the risk of institutionalisation especially the role of receiving home care in delaying institutionalisation and mortality. These findings significantly enhance our

understanding of the prognostic factors and mediators in dementia patients of their adverse prognosis and have implications for their care management. Also, the identification of effective intervention could have valuable implications for service providers and policy makers.

The first contribution of this study is the effect of the home care intervention on dementia patients' institutionalisation and mortality rate and their absolute probability of being discharged to a care institution directly from hospital and death. Dementia pat**ients** with home care had lower rates of hospitalisation after their dementia diagnosis and their subsequent hospitalisations were more likely to be from their usual place of living instead of a care institution. They showed a 6% higher rehospitalisation from outside of care institutions which shows that they were spending more time in their usual place of living. People with home care were spending less time in hospital as the rate of discharge to their usual place of living was 29% higher. This is very important and can have several implications such as shorter stays in hospital for those with home care, which indicates that these dementia patients receive appropriate care which addresses their needs and could also means that their dementia symptoms were not severe enough to be managed in an institutional setting. Also, because of shorter stays in hospital, dementia patients with home care also avoid hospital-associated risk which can improve their overall health outcomes.

I found increasing age and male gender were associated with hospital admission after dementia diagnosis. Similarly, increasing age and male gender were associated with increasing rate of mortality inside and outside of hospital and also within institutional care settings. Similar results of increasing mortality with increasing age were found in other studies [202–205]. In another study by Thearneau et al, showed

that death rate of male with dementia was 1.3 that of females [206]. It is evident from these results that age is of course a risk factor for mortality, hospitalisation and institutionalisation and male gender are associated with increased risk of death with and without a diagnosis of dementia. I also included socioeconomic deprivation as a covariate in the model as it has been identified as a predictor for institutionalisation in the literature [179–181]. I found that increasing socioeconomic deprivation was associated with the risk of hospitalisation after dementia diagnosis and was associated with statistically significant higher risk of subsequent hospital admissions once discharged out of hospital or to a care institution. It was also associated with a higher rate of discharge to care institution directly from hospital. Our results also showed a statistically significant association between socioeconomic deprivation and mortality inside the hospital [179–181]. I found that increasing socioeconomic deprivation was associated with the risk of hospitalisation after dementia diagnosis and was associated with statistically significant higher risk of subsequent hospital admissions once discharged out of hospital or to a care institution. It was also associated with a higher rate of discharge to care institution directly from hospital. The results also showed a statistically significant association between socioeconomic deprivation and mortality inside the hospital. People living in the most deprived vs least deprived areas had a 23% greater hazard of dying in hospital. Also, the hazard of death inside a care institution was reduced with increasing deprivation level, which could be the result of increasing rate of hospitalisation and eventually dying in hospital. These results are consistent with the literature where studies have shown the link between socioeconomic deprivation and poor health outcomes such as cardiovascular

diseases, cognitive impairment, cancer and death [207,208].

Dementia is a syndrome with several subtypes with distinct aetiologies, clinical profiles and outcomes and also there is variation in capability to live across the different subtypes [209]. Therefore, I also looked at the effect of different dementia subtypes on hospitalisations, institutionalisation, and mortality as this has not been looked at before. I found that once patients were diagnosed with dementia, the hospitalisation rate was 14% lower for the AD subtype , whereas it was higher for rare and vascular dementia with a 20% and 6% higher rate of hospitalisation respectively compared to the unspecified dementia type. Similarly, subsequent hospital admissions after discharge after the index hospital admission were only statistically significant from the care institutions. Hospital admissions where the admission source was a long-term care institution were statistically significant for Alzheimer's dementia (AD) and vascular dementia (VaD). The admission rate from care institution was 20% and 41% higher for AD and VaD respectively. This higher rate of hospital admission from care institutions for AD and VaD is explained by their higher rate of discharge to care institutions directly from the hospital and, therefore, they were more likely to be admitted again from care institutions. This could also mean that patients with AD and VaD had severe symptoms and were more likely to be institutionalised because of inadequate support at home to enable independent living. Focusing on the specific needs of AD and VaD patients and providing support to their families and caregivers tailored to their needs can greatly improve their quality of life and decrease the burden on healthcare utilization.

Majority of dementia patients are on cardiovascular drugs and cardiovascular diseases are also a risk factor for dementia. I found that dementia patients who were exposed to anti-cardiovascular drugs before diagnosed with dementia had higher hospitalisation rates after their diagnosis and subsequent hospital admission rates. These patients had a lower rate

of being discharge to a long-term care institution directly from hospital and therefore spending more time in hospital and as a result they were also more likely to die inside hospital which was evident from their higher rate of mortality inside the hospital with a 28% higher hazard ratio of death inside hospital. Also, their death rate in a care institution was also statistically significantly higher (22% higher) compared to those with no cardiovascular complications before dementia diagnosis. Similar to cardiovascular complications, diabetes is also one of the most common comorbid conditions within dementia population as was evident from our literature review and exploratory analysis of our EHR data and therefore exposure to anti-diabetic medication prior to dementia diagnosis and its association with institutionalisation, rehospitalisation and mortality was examined. The results suggested an 20% higher rate of hospitalisation post-dementia diagnosis for those with diabetes compared to those who were not taking any diabetic medications. Dementia patients with diabetes also had higher rate of rehospitalisation both from their usual place of living and from a long-term care institution with an 11% and 29% higher rate respectively. Patients with diabetes were more likely to be discharged from hospital to their usual place of living (6% higher) and the proportion of patient with the availability of home care was higher for those with diabetes (23% vs 21%). This means that they were spending less time in hospital and were going to their own place of living directly from hospital and as a result were more likely to enter care institution from the community where they were living rather than directly from hospital. Their rate of institutionalisation from their usual place of living was 30% higher than non-diabetic dementia patients. In chapter 4, I saw that the risk of dementia diagnosis increased with increasing diabetes duration, and it is also evident here that patients who were diabetic were diagnosed with dementia at a younger age and therefore, timely diabetes diagnosis and effective treatment can delay dementia diagnosis.

Frailty is common in older people and especially dementia patients as they are more

vulnerable to physiological decomposition and studies has shown that frailty is associated with institutionalisation, resource utilisation and mortality[210–214]. I found that with increasing frailty, hospital readmission rate also increased. Similarly, the rate for discharge to long-term care institutions also increased with increasing frailty and dementia patients who were severely frail had a 65% higher rate of institutionalisation directly from hospital and were at an 85% higher rate of institutionalisation from their usual place of living. Increasing frailty was also associated with higher rate of mortality inside and outside of hospital. Patients with severe frailty were elderly with higher rate of hospital readmission and mortality rates. This shows the relationship between the loss of independence with increasing frailty and hospitalisations which results in outcomes such as institutionalisation and mortality.

I also used 30-day readmission as time-varying covariate to check if dementia patients who were re-hospitalised within 30 days after their first hospital discharge had a different rate of transitions to the institutionalisation or death states. In the literature it has been reported that decreasing the length of stay below 10 days leads to an increase in the readmission rate during the 30 days after discharge and increase in the length of stay for some patients may improve their quality of life by reducing readmission during the 30 days after discharge [215,216]. I discovered that individuals with dementia who were readmitted to the hospital within 30 days of their initial discharge had a heightened risk of being hospitalized again, this time from their usual place of residence or a long-term care institution, compared to those who were not readmitted within 30 days. These results are interesting because these patients had a higher rate of transition from hospital to the out state (shortening length of stay) which could be due to the hospital efforts to decrease the hospital length of stay, however, this increases their hospital subsequent readmission both from the Out and Institutionalisation state . These results are similar to those reported in the literature . Therefore, it is very important that a multi-skilled team provide a

comprehensive assessment of the patients physical and psychological needs and there should be good communication between clinicians and taking the patient carer in the decision-making in planning of the patient discharge. This will help to discharge patient to the appropriate setting based on their needs and will reduce long hospital stays but also help those patients who needed to be in the hospital for longer due to their care requirements. The effect of long hospital stays (>= 12 days) and hospital admission due to an injury were also looked at and as these are common outcomes in dementia patients and these outcomes can lead to other severe outcomes such as hospital mortality and discharge to institutional care. Long hospital stays and admission due to injury can act as mediators and explains further the differences in transition to institutionalisation and death between dementia patients. I found that dementia patients who hospital admission was due to an injury had a higher rate of direct discharge from hospital to a care institution and were also more likely to die inside hospital. Similarly, those with hospital stays lasting 12 or more days were more likely to be discharged to long-term care institution and were also an increased risk of death inside hospital because they were spending more time in hospital. Long hospital stays are associated with negative outcomes and as I discussed above that good communication between clinicians and people involved in the care of dementia patients in the decision making to understand the patients' needs will greatly improve the patients' discharge to the appropriate settings.

6.9 Strengths and Limitation

In this study the data comes from the large-scale linked primary care data and hospital episode statistics data which gives an advantage over studies using survey or specialised dementia clinics data because in survey data the number of participants will be small due to the low prevalence of dementia and the data from a specific hospital or specialised clinic suffer from limited variation. In contrast, our data include all participating NHS

practices and the participants hospitalisation data, containing patients with different socioeconomic backgrounds and various health conditions. However, despite the merits of the large electronic data representative of the overall population, there are limitations. First one is that neither the primary care or the hospital episode statistics data are linked to social care datasets containing information on whether the patient had access to informal care and other types of formal care which is not provided by local authorities. The availability of informal care or other types of forma care is associated with homecare usage[195,217]. Informtion about the availability of formal and informal care was derived from the clinical code indicting the provision of formal or informal carer or being a dementia patient and living alone which was taken as proxy for the availability of informal care. This proxy measure is not ideal because it does not capture informal care provided by patient's relatives or family members outside of patient's household. The results showed the effect of homecare availability on institutionalisation and death, however, because the data is hospital admissions and discharge data and therefore, and information on homecare availability is not there and therefore, this hinders the effort to directly examine the effect of homecare on institutionalisation for dementia patients. Therefore, further research in needed with appropriate data to investigate the difference within dementia patients and their institutionalisation and mortality factors.

Chapter 7: Main insights, limitations, and potential areas of future research

The overall aim of the thesis was to explore dementia patients' outcomes reported in the published literature and then using the English electronic health records data explore the characteristics of dementia patients, identify risk factors for dementia diagnoses and explore the fatal and not-fatal outcomes and factors associated with these outcomes using an appropriate model to capture all the nuances of the data at hand.

The aim of this final chapter is to summarise the main insights from the different chapters of this thesis, highlight the important lessons learned and any limitations.

7.1 Dementia as a syndrome

The thesis described why dementia is global health priority and its impact on healthcare resources. In chapter 1, I described dementia as a syndrome and highlighted the symptoms, pathology, and the guidelines for diagnosis. I also highlighted the challenges in the development of new drugs, the failures in clinical trials and the potential treatment targets for future research and development.

I learned from chapter two that dementia is a devastating neurodegenerative disease with profound impact on individuals, families, and healthcare systems. I described the types of dementia and its symptoms and the cognitive tests and biomarkers for diagnosing dementia. I learned the challenges for developing disease-modifying therapies and some of the promising avenues.

Amyloid Beta (Aβ) Plaques: Amyloid beta plaques, a hallmark of Alzheimer's disease, are being targeted by numerous trials. Some compounds, like "3K3A-APA," have shown

promise in genetically modified mice by reducing amyloid production. Another potential treatment is the therapeutic antibody PMN310, which is under preclinical development and specifically targets amyloid beta. Nevertheless, many phase II/III trials of promising drugs targeting amyloid have faced challenges, including failures in clinical testing.

The Immune System: The immune system's role in triggering brain inflammation in dementia has led to the exploration of anti-inflammatory drugs as potential treatments. While some drugs like Pioglitazone and Etanercept have shown potential in phase II trials, their phase III trials were terminated. This avenue continues to be explored for its potential to address the inflammatory component of dementia.

Tau Tangles: Tau tangles, another critical aspect of Alzheimer's pathology, are also being targeted. Approaches include inhibiting protein kinases responsible for tau hyperphosphorylation and developing tau acetylation inhibitors. Additionally, certain nonsteroidal anti-inflammatory drugs (NSAIDs) have demonstrated effectiveness against tau-acetylation, showing promise in animal models.

The challenges in developing new dementia treatments are substantial. The blood-brain barrier often limits drug access, and dementia's diverse pathologies make developing a single treatment challenging. The lack of widely accepted surrogate biomarkers poses difficulties in assessing drug efficacy, and the complexity of dementia progression hampers clinical trial designs. Despite these hurdles, research into potential treatments remains critical due to the projected increase in dementia prevalence and the urgent need for effective interventions. Recently, the FDA approved Lacenemab which works by reducing the Amyloid plaques in the brain for early Alzheimer's disease [218].

7.2 Insight gained from the review of literature.

In chapter 2, I conducted a review of literature in a systematic way to characterise longitudinal cohort studies focused on measuring dementia outcomes post diagnosis and identify opportunities to better design future studies and trials. The results of the review were also considered when analysing our electronic record data.

7.2.1 Main insights from the literature review

From the 100 studies included in the review involving over 2 million individuals, most studies had a small sample size (57% N < 500) , short follow-up time (66% <= 3 years) and were predominantly diagnosed with Alzheimer's dementia (85% Alzheimer's dementia patients).

The review provided some valuable insights into the current state of understanding of dementia progression and identified key areas for improvement in research methodologies. Some of the insights gained from the review are.

Lack of comprehensive studies:

Our review revealed a shortage of comprehensive studies which could address the multi-dimensional nature of dementia progression. Most studies focused primarily measuring cognitive decline, while other important aspects such as function, behavior and quality of life were less frequented investigated.

Limited Biomarker Data:

The review highlighted a scarcity of studies using biomarkers data at different time points during the follow-up which is very important for tracking disease progression and understanding the underlying biology of the disease.

Sample Size and Diversity:

Most of the studies were small as described above which limits the generalizability of the findings. Also, there was a geographical bias towards North America and Europe, which indicate a need for studies which a diverse population around the world.

Diagnostic Criteria and Dementia Subtypes:

I saw a lack of diversity of dementia subtypes being investigated and there was a dominance of Alzheimer's dementia (AD) in the cohort. This emphasizes the importance of using standardised and subtype-specific diagnostic criteria to accurately represent the spectrum of dementia cases.

Linked EHR's data:

The review underscores the potential of linked electronic health records (EHRs) to provide a comprehensive view of dementia patients clinical symptoms, diagnoses, and resource utilization. While some studies utilized administrative databases, very few incorporated linked EHRs data, indicating the need to leverage these resources for more accurate and detailed analyses.

Biological Targets:

For disease-modifying therapies, we need a deeper understanding of the biological mechanisms driving dementia progression. The review highlights the importance of engaging with biomarkers and other biological indicators to develop effective treatments that target the underlying biological causes of dementia.

Outcomes Measured and Resource Utilization:

The studies were focused on cognitive decline as the primary outcome measure may not fully capture the impact of dementia on patient's lives. A more comprehensive approach that takes into consideration the functional abilities, behavior, and quality of life could provide insights into resource utilisation and allocation for patient care.

Collaborative Efforts:

The review underscores the necessity of multidisciplinary coloration among researchers, clinicians, data scientists, and experts in various fields. Collaborative efforts can lead to more comprehensive study designs and better interpretation of the results.

Global Relevance:

The global prevalence of dementia contrasts with the regional bias in research studies. Insights from this review highlights the need for more studies in diverse regions to account for potential variations in dementia progression across different cultural and healthcare systems.

In summary, the literature review offers insights into the gaps and limitations in current research on dementia progression. It emphasizes the importance of a more holistic approach to studying the disease, incorporating longitudinal biomarkers data, standardised diagnostic criteria, diverse cohort of patients, linked EHRs data, and engaging biological targets. Addressing these insights will contribute to a more comprehensive understanding of dementia progression and inform the development of effective disease modifying therapies and patient-centered care.

7.3 Insights gained from the exploratory analysis of our electronic healthcare data.

In chapter 3, I provided an overview of the data sources used to identify our dementia cohort. I described the data sources, its structure and used these sources to identify dementia patients, provided some summary statistics and examined their comorbidities. The insights gained from the exploratory data analysis are:

1. Data Sources and Structure:

The analysis is based on two primary data sources – the Clinical Practice Research Datalink (CPRD) and the Hospital Episode Statistics (HES). CPRD provides anonymized primary care data from UK GP practices, while HES includes information on admissions, discharges, outpatient appointments, and A&E attendances in the NHS hospitals. Both data sources offer valuable information for studying dementia.

2. CPRD Data Structure:

CPRD data is organized into various tables containing distinct types of information such as patient demographics, clinical diagnoses, medications prescribed, test results, referrals, and more. These tables can be joined using a unique patient identifier and this enables the analysis of multiple aspects of patient's healthcare experiences.

3. HES Data Structure:

Hospital Episode Statistics (HES) data is structured into episodes and spells, with each row representing a single finished episode of care. HES encompasses inpatient data and can be linked to other datasets like A & E records, and outpatient data, enhancing the comprehensiveness of the analysis.

4. Patient Selection Criteria:

Our analyses focused on patient with research-quality data in CPRD and HES. Patients aged 30 or over at the study start date with at least one year of up-to-standard data were

included. The patient selection process was aimed at ensuring the validity of follow-up periods.

5. Dementia Patients Identification:

I identified dementia patients using Read codes in the CPRD and ICD-10 codes in the HES dataset. The number and percentage of dementia patients were reported with their mean age and percentage of male and female within each data source separately and combined. Most of the dementia patients identified were from the HES data.

6. Dementia Subtypes and Comorbidities:

I found that a higher proportion of dementia patients were recorded as unspecified dementia. Alzheimer's dementia type was the next mostly recorded subtype after unspecified dementia, followed by vascular dementia (VaD). Vascular dementia was mainly reported in the HES data. Majority of dementia patient were in the 71-90 age category and mostly consisted of females.

7. ICD codes and Comorbidities:

From looking at the top 30 ICD- 10 codes and the corresponding comorbidities, I found that hypertension was on top of the list which makes sense because these patients over 60 and hypertension is very common in this age group. Hypertension was followed by type-2 diabetes, followed by atrial fibrillation, disease of the circulatory system and urinary tract infections.

In conclusion, the exploratory analysis of the datasets provided an insight of the dementia population in terms of their proportion in each dataset, their mean age and proportion of men and women. The analysis shed light on the prevalence of dementia patients, their characteristics, and potential associations with other health conditions.

7.4 Insights gained from the Time-to-Event analysis.

In chapter 4, I presented a study on time-to-event analysis in relation to type-2 diabetes and dementia. Diabetes patients were followed until the first diagnosis of dementia. Our review of literature and exploratory analysis of the datasets showed diabetes as the top comorbidity in dementia. The aim was to see if there is any association between diabetes and the risk of dementia diagnosis.

The main insights from the time-to-event analysis were.

The overall incidence of dementia in diabetic patients over 40 between 2008 and 2017 was 18.9 cases per 1000-person years.

Increasing age, female gender, certain ethnicities, widowed/separated marital status, and certain geographic regions in England were associated with higher dementia risk.

Longer diabetes duration correlated with higher dementia risk, while insulin and anti-diabetic medications were linked to lower risk.

Cerebrovascular complications and hip fracture were significantly associated with dementia risk.

Within the 10 years follow-up dementia developed in 13.85% of diabetic patients aged 40 and over who had diabetes from 1-2 years to 6+ years duration.

This analysis provided us some preliminary insights into the risk factors for dementia in diabetic patients. Further research is needed to explore the underlying risk factors and implications identified in this study and factor in things such as diabetic severity and other data such as BMI, genetic factors to further investigate and explore the potential risk factors.

Insights learned from the concepts in multi-state modelling in chapter 4 First one is that neither the primary care or the hospital episode statistics data are linked to social care datasets containing information on whether the patient had access to informal care and other types of formal care which is not provided by local authorities. The availability of informal care or other types of forma care is associated with homecare usage[195,217]. Informtion about the availability of formal and informal care was derived from the clinical code indicting the provision of formal or informal carer or being a dementia patient and living alone which was taken as proxy for the availability of informal care. This proxy measure is not ideal because it does not capture informal care provided by patient's relatives or family members outside of patient's household. The results showed the effect of homecare availability on institutionalisation and death, however, because the data is hospital admissions and discharge data and therefore, and information on homecare availability is not there and therefore, this hinders the effort to directly examine the effect of homecare on institutionalisation for dementia patients. Therefore, further research in needed with appropriate data to investigate the difference within dementia patients and their institutionalisation and mortality factors.

7.5 Insights gained from the Multi-state models

In chapter 5, I discussed some concepts related to multi-state models in the context of survival analysis and its application to understanding transitions between different states of an illness or disease pathway. Here are the insights that can learned from this chapter, and it gave us a good understanding of why multi-state model are appropriate when we have data on repeat events and provided a good starting point for the next chapter in which I used multi-state model for our data to study dementia patient's hospitalisation, institutionalisation and mortality.

The insights learned are below.

1. Traditional survival analysis and Cox model:

   The traditional survival analysis focuses on two states: alive and dead. A Cox proportional hazard model is used to understand the factors influencing the hazard of a particular outcome such as in chapter 5 I used a Cox model to understand the factors associated with dementia diagnosis in diabetic patients.

2. Multi-state Models:

   Multi-state models extend survival analysis to situations where there are multiple intermediate states (transient states) between the initial and final state (absorbing state). These models are useful when individuals can transition between various states over time, such as hospitalisation, discharge home, and institutionalization.

3. Holistic Approach:

   Unlike traditional Cox models, multi-state models take a holistic approach. They consider all transitions between states, enabling the modelling of individual's progression through various states in a disease pathway.

4. Transition Probabilities and Intensities:

   Multi-state processes are characterized by transition probabilities or transition intensities or hazard. Transition probabilities represent the likelihood of moving from one state to another at a specific time, given the history up to that point. Transition intensities represent the instantaneous hazard of moving to a new state, i.e., how quickly one is going to a state in instantaneous time.

5. Different Model Assumptions:

   - Multi-state models have different assumptions about the dependence of transition rates on time:
   - Time-homogenous Models: Transition intensities remain constant over time.

- Markov Models: Transitions depends only on the current state, indicating no memory of earlier cycles.
- Semi-Markov Models: Duration times depend on the history of the process and time since entry into the state.

6. Stratified Baseline Hazard:

   Multi-state models use stratified baseline hazards for each possible transition, allowing for differences in transition probabilities between states.

7. Transition-specific Covariate Effects:

   Multi-state models consider transition-specific covariate effects, enabling the assessment of different covariate effects on different transitions.

8. Transition Probabilities and their estimation:

   Multi-state models allow estimation of transition probabilities between different states at various time points. The Aalen-Johansen estimator is commonly used for this purpose, and it accounts for the possibility of competing risk.

9. Usefulness of Multi-state models:

   Multi-state models are valuable for handling recurrent events, predicting state probabilities, understanding transition rates, and assessing the impact of covariates on transitions. They provide insights for clinical decision-making, prognosis, and health policy planning. It also allows predictions of transition probabilities based on specific covariate profile at different time points.

In summary, I describe why multi-state model provides a comprehensive framework for analysing complex disease pathways, accounting for multiple transitions and their associations with various factors. These models have practical applications in healthcare research and decision-making, offering insights into the progression of illness and treatment and due to these advantages, I used multi-state model to study dementia

patient's hospitalisation, discharged out of hospital to their usual place of living and direct discharge from hospital into a care institution.

7.6 Insights learned from multi-state model for dementia patient's outcomes.

In chapter 6, I discussed a multi-state model to understand the factors influencing hospital admission, institutionalisation, and mortality in dementia patients. I provided insights related to the methodology, data sources, statistical methods, and study objectives. The key insights learned from chapter 6 are as follows:

1. I found that majority of patients were female (66.8%) and males accounted for 33.2%. The mean age at dementia diagnosis was slightly higher for females (82 years) compared to males (~ 80years).
2. The presence of home care seems to influence several outcomes. Patients with home care have higher rates of hospital discharge to home, lower mortality rates, and decreased transitions to institutional care. This suggests that home care might contribute to better outcomes for dementia patients, including delaying institutionalisation and reducing mortality.
3. Frailty is an important factor influencing transitions between the different states of the model. Higher frailty levels are associated with increased rates of hospitalisation, discharge to home, and discharge to institutional care and death inside and outside of hospital.
4. Increasing age is associated with institutionalisation. Older people are more likely to transition to institutional care from hospital. This suggests that as patient age, they are more likely to require institutional care after being discharged from hospital.

5. Socioeconomic deprivation is identified as a predictor of hospital readmission and institutionalisation. Patients from more deprived areas have higher rates of rehospitalization and institutionalisation, along with increased in-hospital deaths. This highlights the impact of socioeconomic factors on healthcare outcomes for dementia patients.
6. Dementia subtypes and medications: Our results delve into the influence of dementia subtypes and exposure to medications. Patients with Alzheimer's dementia (AD) show different rates of hospital discharge and institutionalisation compared to those with unspecified dementia. The use of anti-cardiovascular and anti-diabetic medications also affects hospitalisation, institutionalisation, and mortality rates.
7. Hospital stays and injury: The length of hospital stays and admission due to an injury play role in patient outcomes. Longer hospital stays are associated with higher rates of discharge to institutional care and increased in-hospital mortality. Admission due to injury leads to higher institutionalisation rates and in-hospital deaths.
8. 30-day readmission: Patients readmitted within 30 days have a higher rates of hospital discharge (shortening length of stay) to home but also higher readmission rates, institutionalisation, and death rates, indicating a complex relationship between readmission and subsequent outcomes.

In summary I learned that this study has implications for care management of dementia patients. Factors such as home care, age, gender, frailty, socioeconomic status, and medications all play significant role in influencing patient outcomes. Understanding these relationships can guide healthcare professionals in providing more tailored care strategies.

7.7 Limitations of the thesis

This thesis, while comprehensive in its scope, has certain limitations that must be acknowledged. Firstly, the literature review only focused on studies with participants diagnosed with dementia aged 65 or over and published in English, based on searching in OVID-MEDLINE alone. The search criteria were limited to time period util 2019 and an updated search may result in more recent studies.

The use of electronic health records (EHR) data is subject to inaccuracies and missing information, which can affect the reliability of the findings. The data sources, primarily the Clinical Practice Research Datalink (CPRD) and the Hospital Episode Statistics (HES), might not capture all relevant patient interactions, leading to potential underreporting of some conditions and outcomes. Additionally, the observational nature of the study means that causality cannot be established between identified risk factors and outcomes. The reliance on specific diagnostic codes to identify dementia patients and their comorbidities might introduce classification bias. . Furthermore, the geographic focus on England may limit the generalizability of the findings to other regions with different healthcare systems and demographic profiles. Therefore, the limitation of this study is that I only used HES data for diagnoses and there was no matching of cohorts done based on age, sex, or diagnosis date. The study also lacks important clinical data such as smoking, alcohol consumption, BMI, genetic information which could be important factors affecting dementia risk.

Another limitation of the data is that neither the primary care or the hospital episode statistics data are linked to social care datasets containing information on whether the patient had access to informal care and other types of formal care which is not provided by local authorities. The availability of informal care or other types of formal

care is associated with homecare usage. Informtion about the availability of formal and informal care was derived from the clinical code indicting the provision of formal or informal carer or being a dementia patient and living alone which was taken as proxy for the availability of informal care. This proxy measure is not ideal because it does not capture informal care provided by patient's relatives or family members outside of patient's household. The results showed the effect of homecare availability on institutionalisation and death, however, because the data is hospital admissions and discharge data and therefore, and information on homecare availability is not there and therefore, this hinders the effort to directly examine the effect of homecare on institutionalisation for dementia patients. Therefore, further research in needed with appropriate data to investigate the difference within dementia patients and their institutionalisation and mortality factors.

7.8 Clinical Implementation and Management of Dementia Patients in the NHS

The findings from this thesis can inform several aspects of clinical practice and patient management within the NHS. Firstly, the identification of key risk factors for dementia, such as diabetes duration and comorbidities like hypertension and cardiovascular diseases, highlights the need for integrated care pathways that address these conditions holistically. The insights on the importance of home care in reducing hospital readmissions and delaying institutionalization suggest that enhancing home care services and support systems could improve patient outcomes. Tailored interventions based on patient frailty and socioeconomic status are crucial, as these factors significantly influence hospitalization and mortality rates. The results also underscore the importance of early and accurate diagnosis using standardized criteria

and biomarkers to ensure appropriate and timely interventions. Additionally, healthcare providers should consider the role of medications, such as anti-cardiovascular and anti-diabetic drugs, in managing dementia patients, potentially adjusting treatment plans to mitigate risks. Additionally, the data supports the need for multidisciplinary collaboration in dementia care, involving clinicians, social workers, and care coordinators to provide comprehensive and patient-centered care.

These results are also interesting because these patients who had a higher rate of transition from hospital to the out state (shortening length of stay) which could be due to the hospital efforts to decrease the hospital length of stay, however, this increases their hospital subsequent readmission both from the Out and Institutionalisation state.

Therefore, it is very important that a multi-skilled team provide a comprehensive assessment of the patients physical and psychological needs and there should be good communication between clinicians and taking the patient carer in the decision-making in planning of the patient discharge. This will help to discharge patient to the appropriate setting based on their needs and will reduce long hospital stays but also help those patients who needed to be in the hospital for longer due to their care requirements. Long hospital stays are associated with negative outcomes and good communication between clinicians and people involved in the care of dementia patients in the decision making to understand the patients' needs will greatly improve the patients' discharge to the appropriate settings.

7.9 Future research or what can be done differently.

7.9.1 Addressing Gaps in Home Care Data Availability

The findings from this thesis have highlighted significant associations between home care availability and various outcomes in dementia patients, such as hospital readmissions and delays in institutionalization. However, the use of proxy measures for home care availability due to the lack of direct data poses limitations to the robustness and accuracy of these findings. Future research should aim to overcome these limitations by obtaining and utilizing direct data on home care services. This section outlines a proposed future research study that addresses these gaps.

7.9.1.1 Objectives

1. To obtain and utilize direct data on home care availability and usage among dementia patients.
2. To evaluate the impact of home care availability on hospitalization rates, institutionalization, and overall patient outcomes more accurately.
3. To assess the effectiveness of different types and intensities of home care services in managing dementia.

7.9.1.2 Methodology

7.9.1.2.1 Data Collection

- Primary Data Collection: Conduct surveys and interviews with healthcare providers, caregivers, and patients to gather detailed information on the availability, frequency, and types of home care services.
- Secondary Data Collection: Collaborate with home care agencies and NHS home care services to access administrative records and service utilization data.

- Integration with EHR Data: Link the collected home care data with existing electronic health records (EHR) from sources such as the Clinical Practice Research Datalink (CPRD) and the Hospital Episode Statistics (HES).

7.9.1.2.2 Study Design

- Cohort Study: Design a prospective cohort study involving dementia patients receiving varying levels of home care services. This will involve regular follow-ups to track patient outcomes over time.
- Comparative Analysis: Compare outcomes between patients with access to robust home care services and those with limited or no access to such services. Utilize propensity score matching to control for confounding variables.
- Mixed-Methods Approach: Combine quantitative analysis with qualitative insights from patient and caregiver interviews to understand the nuances of home care effectiveness.

7.9.1.2.3 Outcome Measures

- Primary Outcomes: Hospital readmission rates, time to institutionalization, and mortality rates.
- Secondary Outcomes: Quality of life, caregiver burden, and patient satisfaction with care.

4. Statistical Analysis:

- Multivariate Regression: Use multivariate regression models to assess the relationship between home care availability and patient outcomes, controlling for potential confounders.
- Survival Analysis: Conduct survival analysis to evaluate time-to-event outcomes, such as time to institutionalization and time to death.
- Multi-state modelling approach

- We can use the linked data of primary care, hospital admission and homecare and using the same model in thesis, analyse the rate of transition between the different states of the model between patients with and without homecare availability while adjusting for important covariates such as frailty, age, gender, socioeconomic deprivation and comorbidities.

7.9.1.2.4 Anticipated Challenges

- Data Integration: Ensuring the seamless integration of home care data with existing EHR data may present technical and logistical challenges.
- Data Privacy: Safeguarding patient confidentiality and obtaining necessary consents for data usage will be critical.
- Variability in Home Care Services: Addressing the variability in types and quality of home care services across different providers and regions.

Expected Contributions:

This future research study aims to provide a more accurate and comprehensive understanding of the role of home care in managing dementia. By utilizing direct data on home care availability, the study will address the limitations of proxy measures and offer robust evidence on the effectiveness of home care services. The findings will inform healthcare policies and clinical practices, ultimately improving the quality of care for dementia patients and supporting caregivers.

By proposing this future research study, I aim to build on the insights gained from the current thesis and address the identified data limitations, paving the way for more precise and actionable recommendations in dementia care management.

Additionally, there is a need for more diverse cohort studies that include underrepresented populations from various geographic and ethnic backgrounds to enhance the generalizability of the findings. Further research into the biological

mechanisms underlying dementia, particularly through the use of biomarkers and genetic data, could lead to the discovery of new therapeutic targets and the development of disease-modifying treatments. Finally, exploring the impact of different care models and interventions on patient outcomes through randomized controlled trials would provide robust evidence to inform clinical practice and healthcare policy.

References

1. Winblad B, Amouyel P, Andrieu S, et al. Defeating Alzheimer's disease and other dementias: A priority for European science and society. *Lancet Neurol*. 2016;15(5):455-532. doi:10.1016/S1474-4422(16)00062-4

2. Gaugler J, James B, Johnson T, Scholz K, Weuve J. 2016 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2016;12(4):459-509. doi:10.1016/j.jalz.2016.03.001

3. Prince M, Ali GC, Guerchet M, Prina AM, Albanese E, Wu YT. Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimers Res Ther*. 2016;8:23. doi:10.1186/s13195-016-0188-8

4. World Alzheimer Report 2020 – Design Dignity Dementia: dementia-related design and the built environment, Volume 1.

5. Dementia. Accessed May 6, 2024. https://www.who.int/news-room/fact-sheets/detail/dementia

6. Cummings J, Ritter A, Zhong K. Clinical Trials for Disease-Modifying Therapies in Alzheimer's Disease: A Primer, Lessons Learned, and a Blueprint for the Future. *J Alzheimers Dis JAD*. 2018;64(s1):S3-S22. doi:10.3233/JAD-179901

7. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet Lond Engl*. 2020;396(10248):413-446. doi:10.1016/S0140-6736(20)30367-6

8. Kalaria RN, Maestre GE, Arizaga R, et al. Alzheimer's disease and vascular dementia in developing countries: prevalence, management, and risk factors. *Lancet Neurol*. 2008;7(9):812-826. doi:10.1016/S1474-4422(08)70169-8

9. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. *The Lancet*. 2005;366(9503):2112-2117. doi:10.1016/S0140-6736(05)67889-0

10. Custodio N, Montesinos R, Lira D, Herrera-Pérez E, Bardales Y, Valeriano-Lorenzo L. Mixed dementia: A review of the evidence. *Dement Neuropsychol*. 2017;11(4):364-370. doi:10.1590/1980-57642016dn11-040005

11. Deramecourt V, Slade JY, Oakley AE, et al. Staging and natural history of cerebrovascular pathology in dementia. *Neurology*. 2012;78(14):1043-1050. doi:10.1212/WNL.0b013e31824e8e7f

12. Kalaria RN. Neuropathological diagnosis of vascular cognitive impairment and vascular dementia with implications for Alzheimer's disease. *Acta Neuropathol (Berl)*. 2016;131(5):659-685. doi:10.1007/s00401-016-1571-z

13. Robinson L, Tang E, Taylor JP. Dementia: Timely diagnosis and early intervention. *BMJ Online*. Published online 2015. doi:10.1136/bmj.h3029

14. Alzheimer, Society. *Dementia UK Second Edition - Overview*.; 2007.

15. Iliffe S, Wilcock J, Austin T, et al. Dementia Diagnosis and Management in Primary Care: Developing and testing educational models. *Dementia*. 2002;1(1):11-23. doi:10.1177/147130120200100111

16. Brodaty H, Pond D, Kemp NM, et al. The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc*. 2002;50(3):530-534.

17. Brooke P, Bullock R. Validation of a 6 item cognitive impairment test with a view to primary care usage. *Int J Geriatr Psychiatry*. 1999;14(11):936-940.

18. Overview | Dementia: assessment, management and support for people living with dementia and their carers | Guidance | NICE.

19. Sommerlad A, Perera G, Singh-Manoux A, Lewis G, Stewart R, Livingston G. Accuracy of general hospital dementia diagnoses in England: Sensitivity, specificity, and predictors of diagnostic accuracy 2008-2016. *Alzheimers Dement*. 2018;14(7):933-943. doi:10.1016/j.jalz.2018.02.012

20. Information for people with dementia and their carers.

21. 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 JAD*. 2018;61(4):1301-1310. doi:10.3233/JAD-170572

22. Dahl A, Berg S, Nilsson SE. Identification of dementia in epidemiological research: a study on the usefulness of various data sources. *Aging Clin Exp Res*. 2007;19(5):381-389. doi:10.1007/BF03324718

23. Lashley T, Schott JM, Weston P, et al. Molecular biomarkers of Alzheimer's disease: progress and prospects. Published online 2018. doi:10.1242/dmm.031781

24. Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*. 2014;13(6):614-629. doi:10.1016/S1474-4422(14)70090-0

25. Biagioni MC, Galvin JE. Using biomarkers to improve detection of Alzheimer's disease. *Neurodegener Dis Manag*. 2011;1(2):127-139. doi:10.2217/NMT.11.11

26. Jack CRJ, Bennett DA, Blennow K, et al. A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*. 2016;87(5):539-547. doi:10.1212/WNL.0000000000002923

27. Blennow K, Mattsson N, Scholl M, Hansson O, Zetterberg H. Amyloid biomarkers in Alzheimer's disease. *Trends Pharmacol Sci*. 2015;36(5):297-309. doi:10.1016/j.tips.2015.03.002

28. Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL, Beyreuther K. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc Natl Acad Sci U S A*. 1985;82(12):4245-4249. doi:10.1073/PNAS.82.12.4245

29. Olsson B, Lautner R, Andreasson U, et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol*. 2016;15(7):673-684. doi:10.1016/S1474-4422(16)00070-3

30. Kuhlmann J, Andreasson U, Pannee J, et al. CSF Aβ1–42 – an excellent but complicated Alzheimer's biomarker – a route to standardisation. *Clin Chim Acta*. 2017;467:27-33. doi:10.1016/J.CCA.2016.05.014

31. Formichi P, Battisti C, Radi E, Federico A. Cerebrospinal fluid tau, A beta, and phosphorylated tau protein for the diagnosis of Alzheimer's disease. *J Cell Physiol*. 2006;208(1):39-46. doi:10.1002/jcp.20602

32. Klunk WE, Engler H, Nordberg A, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol*. 2004;55(3):306-319. doi:10.1002/ana.20009

33. Zetterberg H. Review: Tau in biofluids - relation to pathology, imaging and clinical features. *Neuropathol Appl Neurobiol*. 2017;43(3):194-199. doi:10.1111/nan.12378

34. Blennow K, Wallin A, Agren H, Spenger C, Siegfried J, Vanmechelen E. Tau protein in cerebrospinal fluid: a biochemical marker for axonal degeneration in Alzheimer disease? *Mol Chem Neuropathol*. 1995;26(3):231-245. doi:10.1007/BF02815140

35. Hampel H, Buerger K, Zinkowski R, et al. Measurement of phosphorylated tau epitopes in the differential diagnosis of Alzheimer disease: a comparative cerebrospinal fluid study. *Arch Gen Psychiatry*. 2004;61(1):95-102. doi:10.1001/archpsyc.61.1.95

36. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc*. 2018;14(4):535-562. doi:10.1016/j.jalz.2018.02.018

37. Hostetler ED, Walji AM, Zeng Z, et al. Preclinical Characterization of 18F-MK-6240, a Promising PET Tracer for In Vivo Quantification of Human Neurofibrillary Tangles. *J Nucl Med Off Publ Soc Nucl Med*. 2016;57(10):1599-1606. doi:10.2967/jnumed.115.171678

38. Marquie M, Normandin MD, Vanderburg CR, et al. Validating novel tau positron emission tomography tracer [F-18]-AV-1451 (T807) on postmortem brain tissue. *Ann Neurol*. 2015;78(5):787-800. doi:10.1002/ana.24517

39. Scholl M, Lockhart SN, Schonhaut DR, et al. PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron*. 2016;89(5):971-982. doi:10.1016/j.neuron.2016.01.028

40. Brier MR, Gordon B, Friedrichsen K, et al. Tau and Abeta imaging, CSF measures, and cognition in Alzheimer's disease. *Sci Transl Med*. 2016;8(338):338ra66. doi:10.1126/scitranslmed.aaf2362

41. Cho H, Choi JY, Hwang MS, et al. In vivo cortical spreading pattern of tau and amyloid in the Alzheimer disease spectrum. *Ann Neurol*. 2016;80(2):247-258. doi:10.1002/ana.24711

42. Wirth M, Madison CM, Rabinovici GD, Oh H, Landau SM, Jagust WJ. Alzheimer's disease neurodegenerative biomarkers are associated with decreased cognitive function but not beta-amyloid in cognitively normal older individuals. *J*

Neurosci Off J Soc Neurosci. 2013;33(13):5553-5563. doi:10.1523/JNEUROSCI.4409-12.2013

43. Hesse C, Rosengren L, Andreasen N, et al. Transient increase in total tau but not phospho-tau in human cerebrospinal fluid after acute stroke. *Neurosci Lett*. 2001;297(3):187-190.

44. Hesse C, Rosengren L, Andreasen N, et al. Transient increase in total tau but not phospho-tau in human cerebrospinal fluid after acute stroke. *Neurosci Lett*. 2001;297(3):187-190.

45. Ost M, Nylen K, Csajbok L, et al. Initial CSF total tau correlates with 1-year outcome in patients with traumatic brain injury. *Neurology*. 2006;67(9):1600-1604. doi:10.1212/01.wnl.0000242732.06714.0f

46. Skillback T, Rosen C, Asztely F, Mattsson N, Blennow K, Zetterberg H. Diagnostic performance of cerebrospinal fluid total tau and phosphorylated tau in Creutzfeldt-Jakob disease: results from the Swedish Mortality Registry. *JAMA Neurol*. 2014;71(4):476-483. doi:https://dx.doi.org/10.1001/jamaneurol.2013.6455

47. Riemenschneider M, Wagenpfeil S, Vanderstichele H, et al. Phospho-tau/total tau ratio in cerebrospinal fluid discriminates Creutzfeldt-Jakob disease from other dementias. *Mol Psychiatry*. 2003;8(3):343-347. doi:10.1038/sj.mp.4001220

48. Wirth M, Villeneuve S, Haase CM, et al. Associations between Alzheimer disease biomarkers, neurodegeneration, and cognition in cognitively normal older people. *JAMA Neurol*. 2013;70(12):1512-1519. doi:10.1001/jamaneurol.2013.4013

49. Wisse LEM, Butala N, Das SR, et al. Suspected non-AD pathology in mild cognitive impairment. *Neurobiol Aging*. 2015;36(12):3152-3162. doi:10.1016/j.neurobiolaging.2015.08.029

50. Payoux P, Salabert AS. New PET markers for the diagnosis of dementia. *Curr Opin Neurol*. 2017;30(6):608-616. doi:10.1097/WCO.0000000000000489

51. Waldemar G, Burns A. Chapter: Diagnosing Alzheimer's disease in clinical practice. In: Waldemar G, Burns A, eds. 2 ed. Oxford University Press; 2017. doi:10.1093/med/9780198779803.001.0001

52. *Diagnostic Criteria for Dementia*.

53. Jack CRJ, Dickson DW, Parisi JE, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology*. 2002;58(5):750-757.

54. Gosche KM, Mortimer JA, Smith CD, Markesbery WR, Snowdon DA. Hippocampal volume as an index of Alzheimer neuropathology: findings from the Nun Study. *Neurology*. 2002;58(10):1476-1482.

55. Robert P, Leone E, Amieva H, Renaud D. Managing behavioural and psychological symptoms in Alzheimer's disease. In: Waldemar G, Burns A, eds. *Alzheimer's Disease (Oxford Neurology Library)*. Oxford University Press; 2017.

56. Cummings J, Lee G, Ritter A, Zhong K. Alzheimer's disease drug development pipeline: 2018. *Alzheimers Dement N Y N*. 2018;4:195-214. doi:10.1016/j.trci.2018.03.009

57. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res Ther*. 2014;6(4):37. doi:10.1186/alzrt269

58. Cummings J, Reiber C, Kumar P. The price of progress: Funding and financing Alzheimer's disease drug development. *Alzheimers Dement N Y N*. 2018;4:330-343. doi:10.1016/j.trci.2018.04.008

59. Karran E, Hardy J. Antiamyloid therapy for Alzheimer's disease--are we on the right road? *N Engl J Med*. 2014;370(4):377-378. doi:10.1056/NEJMe1313943

60. Alzheimer's Drug Fails in Another Crushing Disappointment - Scientific American. Accessed March 18, 2019. https://www.scientificamerican.com/article/alzheimer-rsquo-s-drug-fails-in-another-crushing-disappointment/

61. Newman Tim. Experimental drug may prevent Alzheimer's disease. Published 2019. Accessed March 16, 2019. https://www.medicalnewstoday.com/articles/324172.php

62. Lazic D, Sagare AP, Nikolakopoulou AM, Griffin JH, Vassar R, Zlokovic B V. 3K3A-activated protein C blocks amyloidogenic BACE1 pathway and improves functional outcome in mice. *J Exp Med*. Published online January 15, 2019:jem.20181035. doi:10.1084/jem.20181035

63. Promising DMT candidate for Alzheimer's could reach trials in 2019. Accessed March 18, 2019. https://www.pharmaceutical-technology.com/comment/promising-dmt-candidatelzheimers-disease-could-reach-trials-in-2019/

64. Heppner FL, Ransohoff RM, Becher B. Immune attack: the role of inflammation in Alzheimer disease. *Nat Rev Neurosci*. 2015;16(6):358-372. doi:10.1038/nrn3880

65. Heneka MT, Carson MJ, El Khoury J, et al. Neuroinflammation in Alzheimer's disease. *Lancet Neurol*. 2015;14(4):388-405. doi:10.1016/S1474-4422(15)70016-5

66. Takeda and Zinfandel Pharmaceuticals Discontinue TOMMORROW Trial Following Planned Futility Analysis. Accessed March 18, 2019. https://www.takeda.com/newsroom/newsreleases/2018/takeda-tommorrow-trial/

67. Butchart J, Brook L, Hopkins V, et al. Etanercept in Alzheimer disease: A randomized, placebo-controlled, double-blind, phase 2 trial. *Neurology*. 2015;84(21):2161-2168. doi:10.1212/WNL.0000000000001617

68. Medina M. An Overview on the Clinical Development of Tau-Based Therapeutics. *Int J Mol Sci*. 2018;19(4). doi:10.3390/ijms19041160

69. Spillantini MG, Goedert M. Tau pathology and neurodegeneration. *Lancet Neurol*. 2013;12(6):609-622. doi:10.1016/S1474-4422(13)70090-5

70. Schneider A, Mandelkow E. Tau-based treatment strategies in neurodegenerative diseases. *Neurother J Am Soc Exp Neurother*. 2008;5(3):443-457. doi:10.1016/j.nurt.2008.05.006

71. Hanger DP, Anderton BH, Noble W. Tau phosphorylation: the therapeutic challenge for neurodegenerative disease. *Trends Mol Med*. 2009;15(3):112-119. doi:10.1016/j.molmed.2009.01.003

72. Medina M, Hernandez F, Avila J. New Features about Tau Function and Dysfunction. *Biomolecules*. 2016;6(2). doi:10.3390/biom6020021

73. Min SW, Chen X, Tracy TE, et al. Critical role of acetylation in tau-mediated neurodegeneration and cognitive deficits. *Nat Med*. 2015;21(10):1154-1162. doi:10.1038/nm.3951

74. Morgan P, Brown DG, Lennard S, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat Rev Drug Discov*. 2018;17(3):167-181. doi:10.1038/nrd.2017.244

75. Gauthier S, Albert M, Fox N, et al. Why has therapy development for dementia failed in the last two decades? *Alzheimers Dement J Alzheimers Assoc*. 2016;12(1):60-64. doi:10.1016/j.jalz.2015.12.003

76. Hampel H, Lista S, Teipel SJ, et al. Perspective on future role of biological markers in clinical therapy trials of Alzheimer's disease: a long-range point of view beyond 2020. *Biochem Pharmacol*. 2014;88(4):426-449. doi:10.1016/j.bcp.2013.11.009

77. Vellas B, Andrieu S, Sampaio C, Wilcock G. Disease-modifying trials in Alzheimer's disease: a European task force consensus. *Lancet Neurol*. 2007;6(1):56-62. doi:10.1016/S1474-4422(06)70677-9

78. Karran E, Mercken M, De Strooper B. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat Rev Drug Discov*. 2011;10(9):698-712. doi:10.1038/nrd3505

79. Carrillo MC, Brashear HR, Logovinsky V, et al. Can we prevent Alzheimer's disease? Secondary "prevention" trials in Alzheimer's disease. *Alzheimers Dement*. 2013;9(2):123-131.e1. doi:10.1016/j.jalz.2012.12.004

80. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc*. 2007;3(3):186-191. doi:10.1016/j.jalz.2007.04.381

81. Hewes L. Titles of Theses. *Prof Geogr*. 1950;2(4):14-20. doi:10.1111/j.0033-0124.1950.24_14.x

82. NICE. Dementia: assessment, management and support for people living with dementia and their carers. *NICE Guidel NG97*. 2018;(June).

83. *RESEARCHING ALZHEIMER'S MEDICINES SETBACKS AND STEPPING STONES*.

84. Miller JB, Cummings J, Nance C, Ritter A. Neuroscience learning from longitudinal cohort studies of Alzheimer's disease: Lessons for disease-modifying drug programs and an introduction to the Center for Neurodegeneration and Translational Neuroscience. *Alzheimers Dement Transl Res Clin Interv*. 2018;4:350-356. doi:10.1016/j.trci.2018.06.006

85. Pujades-Rodriguez M, Assi V, Gonzalez-Izquierdo A, et al. The diagnosis, burden and prognosis of dementia: A record-linkage cohort study in England. Forloni G, ed. *PLOS ONE*. 2018;13(6):e0199026. doi:10.1371/journal.pone.0199026

86. Prince M, Knapp M, Guerchet M, et al. *Dementia UK 2014: Update*.; 2014.

87. Mckhann G, Drachman D, Folstein M. views & reviews Clinical diagnosis of Alzheimer ' s disease : *Neurology*. 1984;34(7):939--944. doi:10.1186/alzrt38

88. Bell CC. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision: DSM-IV-TR Quick Reference to the Diagnostic Criteria from DSM-IV-TR. *JAMA J Am Med Assoc*. 2001;285(6):811-812. doi:10.1001/jama.285.6.811

89. Roman GC, Tatemichi TK, Erkinjuntti T, et al. Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*. 1993;43(2):250-260.

90. Jack CR, Albert MS, Knopman DS, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Published online 2011. doi:10.1016/j.jalz.2011.03.004

91. Vemuri P, Wiste HJ, Weigand SD, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. *Neurology*. 2009;73(4):294-301. doi:https://dx.doi.org/10.1212/WNL.0b013e3181af79fb

92. Brys M, Pirraglia E, Rich K, et al. Prediction and longitudinal study of CSF biomarkers in mild cognitive impairment. *Neurobiol Aging*. 2009;30(5):682-690.

93. Morbelli S, Bauckneht M. Amyloid PET Imaging: Standardization and Integration with Other Alzheimer's Disease Biomarkers. In: Perneczky R, ed. *Biomarkers for Alzheimer's Disease Drug Development*. Springer New York; 2018:203-212. doi:10.1007/978-1-4939-7704-8_13

94. Dumurgier J, Hanseeuw BJ, Hatling FB, et al. Alzheimer's Disease Biomarkers and Future Decline in Cognitive Normal Older Adults. *J Alzheimers Dis JAD*. 2017;60(4):1451-1459. doi:https://dx.doi.org/10.3233/JAD-170511

95. Caroli A, Prestia A, Wade S, et al. Alzheimer Disease Biomarkers as Outcome Measures for Clinical Trials in MCI. *Alzheimer Dis Assoc Disord*. 2015;29(2):101-109. doi:https://dx.doi.org/10.1097/WAD.0000000000000071

96. Allen SJ, Waldemar G, Burns A. Chapter: Pathophysiology of Alzheimer's disease Oxford Medicine Online Key points. Published online 2017. doi:10.1093/med/9780198779803.001.0001

97. Zetzsche T, Rujescu D, Hardy J, Hampel H. Advances and perspectives from genetic research: development of biological markers in Alzheimer's disease. *Expert Rev Mol Diagn*. 2010;10(5):667-690. doi:10.1586/erm.10.48

98. Poirier J. Apolipoprotein E in animal models of CNS injury and in Alzheimer's disease. *Trends Neurosci*. 1994;17(12):525-530.

99. Cosentino S, Scarmeas N, Helzner E, et al. APOE epsilon 4 allele predicts faster cognitive decline in mild Alzheimer disease. *Neurology*. 2008;70(19 Pt 2):1842-1849. doi:https://dx.doi.org/10.1212/01.wnl.0000304038.37421.cc

100. Kemppainen NM, Scheinin NM, Koivunen J, et al. Five-year follow-up of 11C-PIB uptake in Alzheimer's disease and MCI. *Eur J Nucl Med Mol Imaging*. 2014;41(2):283-289. doi:10.1007/s00259-013-2562-0

101. Tapiola T, Overmyer M, Lehtovirta M, et al. The level of cerebrospinal fluid tau correlates with neurofibrillary tangles in Alzheimer's disease. *Neuroreport*. 1997;8(18):3961-3963.

102. Tapiola T, Alafuzoff I, Herukka SK, et al. Cerebrospinal fluid {beta}-amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. *Arch Neurol*. 2009;66(3):382-389. doi:10.1001/archneurol.2008.596

103. Blennow K, Hampel H. CSF markers for incipient Alzheimer's disease. *Lancet Neurol*. 2003;2(10):605-613.

104. Popp J, Lewczuk P, Frommann I, et al. Cerebrospinal fluid markers for Alzheimer's disease over the lifespan: effects of age and the APOEepsilon4 genotype. *J Alzheimers Dis JAD*. 2010;22(2):459-468. doi:10.3233/JAD-2010-100561

105. Grimmer T, Tholen S, Yousefi BH, et al. Progression of cerebral amyloid load is associated with the apolipoprotein E epsilon4 genotype in Alzheimer's disease. *Biol Psychiatry*. 2010;68(10):879-884. doi:https://dx.doi.org/10.1016/j.biopsych.2010.05.013

106. Villemagne VL, Burnham S, Bourgeat P, et al. Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol*. 2013;12(4):357-367. doi:https://dx.doi.org/10.1016/S1474-4422(13)70044-9

107. Lleó A, Alcolea D, Martínez-Lage P, et al. Longitudinal cerebrospinal fluid biomarker trajectories along the Alzheimer's disease continuum in the BIOMARKAPD study. *Alzheimers Dement*. Published online 2019:1-12. doi:10.1016/j.jalz.2019.01.015

108. Lewczuk P, Riederer P, … SOT world journal of, 2018 undefined. Cerebrospinal fluid and blood biomarkers for neurodegenerative dementias: an update of the

Consensus of the Task Force on Biological Markers in Psychiatry of the. *Taylor Francis*.

109. Alcolea D, Martínez-Lage P, Neurology PSJ, 2015 undefined. Amyloid precursor protein metabolism and inflammation markers in preclinical Alzheimer disease. *AAN Enterp*.

110. … MQVJ of, 2017 undefined. YKL-40 (Chitinase 3-like I) is expressed in a subset of astrocytes in Alzheimer's disease and other tauopathies. *Jneuroinflammationbiomedcentral* ….

111. Craig-Schapiro R, Perrin R, Roe C, … CXB, 2010 undefined. YKL-40: a novel prognostic fluid biomarker for preclinical Alzheimer's disease. *Elsevier*.

112. Ne Amieva H, Le Goff M, Millet X, et al. Prodromal Alzheimer's Disease: Successive Emergence of the Clinical Symptoms. Published online 2008. doi:10.1002/ana.21509

113. Vellas B, Andrieu S, Sampaio C, Coley N, Wilcock G. Endpoints for trials in Alzheimer's disease: a European task force consensus. *Lancet Neurol*. 2008;7(5):436-450. doi:10.1016/S1474-4422(08)70087-5

114. Sampaio C. Clinical relevance on Alzheimer's disease endpoints. *J Nutr Health Aging*. 2007;11(4):316-317.

115. Cook M, Baker N, Lanes S, Bullock R, Wentworth C, Michael Arrighi H. Incidence of stroke and seizure in Alzheimer's disease dementia. *Age Ageing*. 2015;44(4):695-699. doi:10.1093/ageing/afv061

116. Browne J, Edwards DA, Rhodes KM, Brimicombe DJ, Payne RA. Association of comorbidity and health service usage among patients with dementia in the UK: A population-based study. *BMJ Open*. Published online 2017. doi:10.1136/bmjopen-2016-012546

117. Zhao Y, Kuo TC, Weir S, Kramer MS, Ash AS. Healthcare costs and utilization for Medicare beneficiaries with Alzheimer's. *BMC Health Serv Res*. 2008;8:108. doi:https://dx.doi.org/10.1186/1472-6963-8-108

118. Griffith LE, Gruneir A, Fisher K, et al. Patterns of health service use in community living older adults with dementia and comorbid conditions: a population-based retrospective cohort study in Ontario, Canada. *BMC Geriatr*. Published online 2016. doi:10.1186/s12877-016-0351-x

119. Lopez OL, Swihart AA, Becker JT, et al. Reliability of NINCDS-ADRDA clinical criteria for the diagnosis of Alzheimer's disease. *Neurology*. 1990;40(10):1517-1522.

120. Blacker D, Albert MS, Bassett SS, Go RC, Harrell LE, Folstein MF. Reliability and validity of NINCDS-ADRDA criteria for Alzheimer's disease. The National Institute of Mental Health Genetics Initiative. *Arch Neurol*. 1994;51(12):1198-1204.

121. Baldereschi M, Amato MP, Nencini P, et al. Cross-national interrater agreement on the clinical diagnostic criteria for dementia. WHO-PRA Age-Associated

Dementia Working Group, WHO-Program for Research on Aging, Health of Elderly Program. *Neurology*. 1994;44(2):239-242.

122. Hill J, Fillit H, Shah SN, del Valle MC, Futterman R. Patterns of healthcare utilization and costs for vascular dementia in a community-dwelling population. *J Alzheimers Dis JAD*. 2005;8(1):43-50.

123. Silverman DHS, Gambhir SS, Huang HWC, et al. Evaluating early dementia with and without assessment of regional cerebral metabolism by PET: a comparison of predicted costs and benefits. *J Nucl Med Off Publ Soc Nucl Med*. 2002;43(2):253-266.

124. Magin P, Juratowitch L, Dunbabin J, et al. Attitudes to Alzheimer's disease testing of Australian general practice patients: a cross-sectional questionnaire-based study. *Int J Geriatr Psychiatry*. 2016;31(4):361-366. doi:10.1002/gps.4335

125. Carrillo MC, Brashear HR, Logovinsky V, et al. Can we prevent Alzheimer's disease? Secondary "prevention" trials in Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc*. 2013;9(2):123-131.e1. doi:10.1016/j.jalz.2012.12.004

126. Morgan P, Brown DG, Lennard S, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat Rev Drug Discov*. 2018;17(3):167-181. doi:10.1038/nrd.2017.244

127. Disorders C. NIA commentary on the NIA-AA Research Framework : Towards a biological definition of Alzheimer ' s disease. 2018;14:576-578. doi:10.1016/j.jalz.2018.03.004

128. Cummings J, Aisen PS, DuBois B, et al. Drug development in Alzheimer's disease: the path to 2025. *Alzheimers Res Ther*. 2016;8:39. doi:10.1186/s13195-016-0207-9

129. Lees R, Fearon P, Harrison JK, Broomfield NM, Quinn TJ. Cognitive and Mood Assessment in Stroke Research. *Stroke*. 2012;43(6):1678-1680. doi:10.1161/STROKEAHA.112.653303

130. Fda, Cder, Mccrayk. *Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease*.

131. Browne J, Edwards DA, Rhodes KM, Brimicombe DJ, Payne RA. Association of comorbidity and health service usage among patients with dementia in the UK: a population-based study. *BMJ Open*. 2017;7(3):e012546. doi:https://dx.doi.org/10.1136/bmjopen-2016-012546

132. Chen L, Reed C, Happich M, Nyhuis A, Lenox-Smith A. Health care resource utilisation in primary care prior to and after a diagnosis of Alzheimer's disease: a retrospective, matched case-control study in the United Kingdom. *BMC Geriatr*. 2014;14:76. doi:https://dx.doi.org/10.1186/1471-2318-14-76

133. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836. doi:10.1093/ije/dyv098

134. Padmanabhan S, Carty L, Cameron E, Ghosh RE, Williams R, Strongman H. Approach to record linkage of primary care data from Clinical Practice Research

Datalink to other health-related patient data: overview and implications. *Eur J Epidemiol*. 2019;34(1):91-99. doi:10.1007/s10654-018-0442-4

135. Hospital Episode Statistics (HES) - NHS Digital. Accessed February 24, 2020. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics

136. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol*. 2017;46(4):1093-1093i. doi:10.1093/ije/dyx015

137. Boyd A, Cornish R, Johnson L, et al. *Understanding Hospital Episode Statistics (HES) Resource Report*.; 2018.

138. Bendlin BB. Antidiabetic therapies and Alzheimer disease. *Dialogues Clin Neurosci*. 2019;21(1):83-91.

139. Ott A, Stolk RP, Hofman A, van Harskamp F, Grobbee DE, Breteler MMB. Association of diabetes mellitus and dementia: The Rotterdam Study. *Diabetologia*. 1996;39(11):1392-1397. doi:10.1007/s001250050588

140. R core team. R: A language and environment for statistical computing. Published online 2019.

141. Jaul E, Barron J. Age-Related Diseases and Clinical and Public Health Implications for the 85 Years Old and Over Population. *Front Public Health*. 2017;5:335. doi:10.3389/fpubh.2017.00335

142. Roglic G, World Health Organization. *Global Report on Diabetes*.

143. Ott A, Stolk RP, van Harskamp F, Pols HA, Hofman A, Breteler MM. Diabetes mellitus and the risk of dementia: The Rotterdam Study. *Neurology*. 1999;53(9):1937-1942. doi:10.1212/wnl.53.9.1937

144. Xu W, Caracciolo B, Wang HX, et al. Accelerated Progression From Mild Cognitive Impairment to Dementia in People With Diabetes. *Diabetes*. 2010;59(11):2928-2935. doi:10.2337/db10-0539

145. *Title Dementia UK: Update Second Edition Acknowledgements The Authors Would like to Thank Everyone Involved in The*.

146. Xu W, Caracciolo B, Wang HX, et al. Accelerated Progression From Mild Cognitive Impairment to Dementia in People With Diabetes. *Diabetes*. 2010;59(11):2928-2935. doi:10.2337/db10-0539

147. 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(10):1868-1875. doi:10.2337/dc15-0491

148. Qiu C, De Ronchi D, Fratiglioni L. The epidemiology of the dementias: an update. *Curr Opin Psychiatry*. 2007;20(4):380-385. doi:10.1097/YCO.0b013e32816ebc7b

149. Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Heterogeneity in 14-year Dementia Incidence Between Asian American Subgroups. *Alzheimer Dis Assoc Disord*. 2017;31(3):181-186. doi:10.1097/WAD.0000000000000189

150. Mukadam N, Waugh A, Cooper C, Livingston G. What would encourage help-seeking for memory problems among UK-based South Asians? A qualitative study Setting: Community settings in and around Greater. doi:10.1136/bmjopen-2015-007990

151. Sommerlad A, Ruegger J, Singh-Manoux A, Lewis G, Livingston G. Marriage and risk of dementia: systematic review and meta-analysis of observational studies. *J Neurol Neurosurg Psychiatry*. 2018;89(3):231-238. doi:10.1136/JNNP-2017-316274

152. 2011 Census - Office for National Statistics. 2011. Accessed April 26, 2020. https://www.ons.gov.uk/census/2011census

153. Prevalence of diabetes in England. Published 2018. Accessed April 26, 2020. https://www.diabetes.co.uk/news/2019/oct/the-midlands-has-the-highest-rates-of-diabetes-in-england-97468460.html

154. Baker C. Diabetes in England: where are the hotspots?

155. *The English Indices of Deprivation 2019*.; 2019.

156. Bruce DG, Davis WA, Casey GP, et al. Predictors of cognitive impairment and dementia in older people with diabetes. *Diabetologia*. 2008;51(2):241-248. doi:10.1007/s00125-007-0894-7

157. Beeri MS, Schmeidler J, Silverman JM, et al. Insulin in combination with other diabetes medication is associated with less Alzheimer neuropathology. *Neurology*. 2008;71(10):750-757. doi:10.1212/01.wnl.0000324925.95210.6d

158. Hsu CC, Wahlqvist ML, Lee MS, Tsai HN. Incidence of Dementia is Increased in Type 2 Diabetes and Reduced by the Use of Sulfonylureas and Metformin. *J Alzheimers Dis*. 2011;24(3):485-493. doi:10.3233/JAD-2011-101524

159. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007;26(11):2389-2430. doi:10.1002/sim.2712

160. Therneau T, Crowson C, Atkinson E. Multi-state models and competing risks 1 Multi-state models. 2016;(0):1-27.

161. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94(446):496-509. doi:10.1080/01621459.1999.10474144

162. Meira-machado L, Uña-álvarez J De, Cardarso-suárez C. Multi-state models for the analysis of time to event.

163. Sutradhar R, Forbes S, Urbach DR, Paszat L, Rabeneck L, Baxter NN. Multistate models for comparing trends in hospitalizations among young adult survivors of colorectal cancer and matched controls. *BMC Health Serv Res*. 2012;12(1):1-7. doi:10.1186/1472-6963-12-353

164. Borgan Ø. A alen– J ohansen Estimator . *Wiley StatsRef Stat Ref Online*. Published online 2018:1-13. doi:10.1002/9781118445112.stat06001.pub2

165. Ieva F, Jackson CH, Sharples LD. Multi-state modelling of repeated hospitalisation and death in patients with heart failure: The use of large administrative databases in clinical epidemiology. *Stat Methods Med Res*. 2017;26(3):1350-1372. doi:10.1177/0962280215578777

166. Postmus D, van Veldhuisen DJ, Jaarsma T, et al. The COACH risk engine: a multistate model for predicting survival and hospitalization in patients with heart failure. *Eur J Heart Fail*. 2012;14(2):168-175. doi:10.1093/eurjhf/hfr163

167. Ieva F, Jackson CH, Sharples LD. Multi-state modelling of repeated hospitalisation and death in patients with heart failure: The use of large administrative databases in clinical epidemiology. *Stat Methods Med Res*. 2017;26(3):1350-1372. doi:10.1177/0962280215578777

168. Bakal JA, McAlister FA, Liu W, Ezekowitz JA. Heart failure re-admission: measuring the ever shortening gap between repeat heart failure hospitalizations. *PloS One*. 2014;9(9):e106494. doi:10.1371/journal.pone.0106494

169. Commenges D, Joly P. Multi-state Model for Dementia, Institutionalization, and Death. *Commun Stat - Theory Methods*. 2004;33(6):1315-1326. doi:10.1081/STA-120030151

170. Mjørud M, Selbæk G, Bjertness E, et al. Time from dementia diagnosis to nursing-home admission and death among persons with dementia: A multistate survival analysis. *PloS One*. 2020;15(12):e0243513. doi:10.1371/journal.pone.0243513

171. Jack CR, Therneau TM, Wiste HJ, et al. Transition rates between amyloid and neurodegeneration biomarker states and to dementia: a population-based, longitudinal cohort study. *Lancet Neurol*. 2016;15(1):56-64. doi:10.1016/S1474-4422(15)00323-3

172. Jack CR, Therneau TM, Lundt ES, et al. Long-term associations between amyloid positron emission tomography, sex, apolipoprotein E and incident dementia and mortality among individuals without dementia: hazard ratios and absolute risk. *Brain Commun*. 2022;4(2):fcac017. doi:10.1093/braincomms/fcac017

173. Knox S, Downer B, Haas A, Middleton A, Ottenbacher KJ. Dementia Severity Associated With Increased Risk of Potentially Preventable Readmissions During Home Health Care. *J Am Med Dir Assoc*. 2020;21(4):519-524.e3. doi:10.1016/J.JAMDA.2019.09.012

174. Ma C, Bao S, Dull P, Wu B, Yu F. Hospital readmission in persons with dementia: A systematic review. *Int J Geriatr Psychiatry*. 2019;34(8):1170-1184. doi:10.1002/gps.5140

175. Chappell NL, Dlitt BH, Hollander MJ, Miller JA, McWilliam C. Comparative costs of home care and residential care. *The Gerontologist*. 2004;44(3):389-400. doi:10.1093/geront/44.3.389

176. Hollander MJ, Chappell NL. A comparative analysis of costs to government for home care and long-term residential care services, standardized for client care needs. *Can J Aging Rev Can Vieil*. 2007;26 Suppl 1:149-161. doi:10.3138/cja.26.suppl_1.149

177. *England Safe, Compassionate Care for Frail Older People Using an Integrated Care Pathway*.

178. Eska K, Graessel E, Donath C, Schwarzkopf L, Lauterberg J, Holle R. Predictors of Institutionalization of Dementia Patients in Mild and Moderate Stages: A 4-Year Prospective Analysis. *Dement Geriatr Cogn Disord Extra*. 2013;3(1):426-445. doi:10.1159/000355079

179. Agüero-Torres H, von Strauss E, Viitanen M, Winblad B, Fratiglioni L. Institutionalization in the elderly: the role of chronic diseases and dementia. Cross-sectional and longitudinal data from a population-based study. *J Clin Epidemiol*. 2001;54(8):795-801. doi:10.1016/s0895-4356(00)00371-1

180. Andel R, Hyer K, Slack A. Risk Factors for Nursing Home Placement in Older Adults With and Without Dementia. *J Aging Health*. 2007;19(2):213-228. doi:10.1177/0898264307299359

181. Yaffe K, Fox P, Newcomer R, et al. Patient and caregiver characteristics and nursing home placement in patients with dementia. *JAMA*. 2002;287(16):2090-2097. doi:10.1001/jama.287.16.2090

182. Luppa M, Luck T, Weyerer S, König HH, Brähler E, Riedel-Heller SG. Prediction of institutionalization in the elderly. A systematic review. *Age Ageing*. 2010;39(1):31-38. doi:10.1093/ageing/afp202

183. Lau B, Cole SR, Gange SJ. Competing Risk Regression Models for Epidemiologic Data. *Am J Epidemiol*. 2009;170(2):244-256. doi:10.1093/aje/kwp107

184. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation*. 2016;133(6):601-609. doi:10.1161/CIRCULATIONAHA.115.017719

185. Lumley T, S- R, Elizabeth A, Cynthia C, Therneau MTM. Package ' survival .' Published online 2021.

186. R: The R Project for Statistical Computing. Accessed December 13, 2022. https://www.r-project.org/

187. *Alzheimer's Society Analysis of NHS England's Hospital Episode Statistics Dataset*.; 2012.

188. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing*. 2016;45(3):353-360. doi:10.1093/ageing/afw039

189. Hollander MJ, Chappell NL. A comparative analysis of costs to government for home care and long-term residential care services, standardized for client care needs. *Can J Aging Rev Can Vieil*. 2007;26 Suppl 1(S1):149-161. doi:10.3138/cja.26.suppl_1.149

190. Branch LG, Jette AM. A prospective study of long-term care institutionalization among the aged. *Am J Public Health*. 1982;72(12):1373-1379. doi:10.2105/AJPH.72.12.1373

191. Jette AM, Tennstedt S, Crawford S. How does formal and informal community care affect nursing home use? *J Gerontol - Ser B Psychol Sci Soc Sci*. 1995;50 B(1):S4-S12. doi:10.1093/geronb/50B.1.S4

192. Lo Sasso AT, Johnson RW. Does informal care from adult children reduce nursing home admissions for the elderly? *Inquiry*. 2002;39(3):279-297. doi:10.5034/inquiryjrnl_39.3.279

193. Bonsang E. Does informal care from children to their elderly parents substitute for formal care in Europe? *J Health Econ*. 2009;28(1):143-154. doi:10.1016/j.jhealeco.2008.09.002

194. Van Houtven CH, Norton EC. Informal care and health care use of older adults. *J Health Econ*. 2004;23(6):1159-1180. doi:10.1016/j.jhealeco.2004.04.008

195. Lyons KS, Zarit SH. Formal and informal support: the great divide. *Int J Geriatr Psychiatry*. 1999;14(3):183-186.

196. Kojima G. Frailty as a predictor of hospitalisation among community-dwelling older people: a systematic review and meta-analysis. *J Epidemiol Community Health*. 2016;70(7):722-729. doi:10.1136/jech-2015-206978

197. Haaksma ML, Rizzuto D, Ramakers IHGB, et al. The Impact of Frailty and Comorbidity on Institutionalization and Mortality in Persons With Dementia: A Prospective Cohort Study. *J Am Med Dir Assoc*. 2019;20(2):165-170.e2. doi:10.1016/j.jamda.2018.06.020

198. Chehrehgosha M, Vatan RF, Alizadeh Khoei M, et al. Role of frailty in prediction of hospitalized older adult patient's outcomes: a prospective study. *Turk J Med Sci*. 2021;51(5):2324-2333. doi:10.3906/sag-2012-332

199. de Wreede LC, Fiocco M, Putter H. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. *Comput Methods Programs Biomed*. 2010;99(3):261-274. doi:10.1016/j.cmpb.2010.01.001

200. Putter H, van der Hage J, de Bock GH, Elgalta R, van de Velde CJH. Estimation and Prediction in a Multi-State Model for Breast Cancer. *Biom J*. 2006;48(3):366-380. doi:10.1002/bimj.200510218

201. Hardman RJ, Kennedy G, Macpherson H, et al. Resourse Use and Disease Couse in dementia - Nursing Home (REDIC-NH), a longitudinal cohort study; design and patient characteristics at admission to Norwegian nursing homes. Erkinjuntti T Pinhanen P, Ylikoski R, Jokinen H, Somerkoski MM, Mantyla R, Salonen O, Fazekas F, Schmidt R, Ropele S, Rous B, Petrovic K, Garmehi U, Seewann A, Ferro JM, Verdelho A, Madureira S, Moleiro C, Scheltens P, van Straaten I, Barkhof F, Gouw A, PT, ed. *J Alzheimers Dis JAD*. 2016;29(1):1-8. doi:https://dx.doi.org/10.1371/journal.pone.0049072

202. Joling KJ, Janssen O, Francke AL, et al. Time from diagnosis to institutionalization and death in people with dementia. *Alzheimers Dement J Alzheimers Assoc*. 2020;16(4):662-671. doi:10.1002/alz.12063

203. De Matteis G, Burzo ML, Della Polla DA, et al. Outcomes and Predictors of In-Hospital Mortality among Older Patients with Dementia. *J Clin Med*. 2022;12(1):59. doi:10.3390/jcm12010059

204. Sampson EL, Blanchard MR, Jones L, Tookman A, King M. Dementia in the acute hospital: prospective cohort study of prevalence and mortality. *Br J Psychiatry J Ment Sci*. 2009;195(1):61-66. doi:10.1192/bjp.bp.108.055335

205. Xie J, Brayne C, Matthews FE, Medical Research Council Cognitive Function and Ageing Study collaborators. Survival times in people with dementia: analysis from population based cohort study with 14 year follow-up. *BMJ*. 2008;336(7638):258-262. doi:10.1136/bmj.39433.616678.25

206. Jack CR, Therneau TM, Lundt ES, et al. Long-term associations between amyloid positron emission tomography, sex, apolipoprotein E and incident dementia and mortality among individuals without dementia: hazard ratios and absolute risk. *Brain Commun*. 2022;4(2). doi:10.1093/braincomms/fcac017

207. McCann A, McNulty H, Rigby J, et al. Effect of Area-Level Socioeconomic Deprivation on Risk of Cognitive Dysfunction in Older Adults. *J Am Geriatr Soc*. 2018;66(7):1269-1275. doi:10.1111/jgs.15258

208. Shea S, Lima J, Diez-Roux A, Jorgensen NW, McClelland RL. Socioeconomic Status and Poor Health Outcome at 10 Years of Follow-Up in the Multi-Ethnic Study of Atherosclerosis. *PloS One*. 2016;11(11):e0165651. doi:10.1371/journal.pone.0165651

209. Wu YT, Clare L, Hindle J V., Nelis SM, Martyr A, Matthews FE. Dementia subtype and living well: results from the Improving the experience of Dementia and Enhancing Active Life (IDEAL) study. *BMC Med*. 2018;16(1):140. doi:10.1186/s12916-018-1135-2

210. Rockwood K, Mitnitski A, Song X, Steen B, Skoog I. Long-term risks of death and institutionalization of elderly people in relation to deficit accumulation at age 70. *J Am Geriatr Soc*. 2006;54(6):975-979. doi:10.1111/j.1532-5415.2006.00738.x

211. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56(3):M146-56. doi:10.1093/gerona/56.3.m146

212. Mousa A, Savva GM, Mitnitski A, et al. Is frailty a stable predictor of mortality across time? Evidence from the Cognitive Function and Ageing Studies. *Age Ageing*. 2018;47(5):721-727. doi:10.1093/ageing/afy077

213. Jin HY, Liu X, Xue QL, Chen S, Wu C. The Association between Frailty and Healthcare Expenditure among Chinese Older Adults. *J Am Med Dir Assoc*. 2020;21(6):780-785. doi:10.1016/j.jamda.2020.03.008

214. Puts MTE, Lips P, Ribbe MW, Deeg DJH. The effect of frailty on residential/nursing home admission in the Netherlands independent of chronic diseases and functional limitations. *Eur J Ageing*. 2005;2(4):264-274. doi:10.1007/s10433-005-0011-z

215. Carey K, Lin MY. Hospital length of stay and readmission: an early investigation. *Med Care Res Rev MCRR*. 2014;71(1):99-111. doi:10.1177/1077558713504998

216. Figueroa R, Harman J, Engberg J. Use of claims data to examine the impact of length of inpatient psychiatric stay on readmission rate. *Psychiatr Serv Wash DC*. 2004;55(5):560-565. doi:10.1176/appi.ps.55.5.560

217. Bonsang E. Does informal care from children to their elderly parents substitute for formal care in Europe? *J Health Econ*. 2009;28(1):143-154. doi:10.1016/j.jhealeco.2008.09.002

218. FDA Converts Novel Alzheimer's Disease Treatment to Traditional Approval | FDA. Accessed August 20, 2023. https://www.fda.gov/news-events/press-announcements/fda-converts-novel-alzheimers-disease-treatment-traditional-approval

Appendix A

Table 14 : ICD-10 codes used to identify dementia patients.

ICD	LABEL	AGG
F00	Dementia in Alzheimer's disease	Alzheimer's dementia
F00.0	Dementia in Alzheimer's disease with early onset	Alzheimer's dementia
F00.1	Dementia in Alzheimer's disease with late onset	Alzheimer's dementia
F00.2	Dementia in Alzheimer's disease atypical or mixed type	Alzheimer's dementia
F00.9	Dementia in Alzheimer's disease unspecified	Alzheimer's dementia
G30	Alzheimer's disease	Alzheimer's dementia
G30.0	Alzheimer's disease with early onset	Alzheimer's dementia
G30.1	Alzheimer's disease with late onset	Alzheimer's dementia
G30.8	Other Alzheimer disease unspecified	Alzheimer's dementia
G30.9	Alzheimer's disease unspecified	Alzheimer's dementia
F01	Vascular dementia	Vascular dementia
F01.0	Vascular dementia of acute onset	Vascular dementia
F01.1	Multi-infarct dementia	Vascular dementia
F01.2	Subcortical vascular dementia	Vascular dementia
F01.3	Mixed cortical and subcortical vascular dementia	Vascular dementia
F01.8	Other vascular dementia	Vascular dementia
F01.9	Vascular dementia, unspecified	Vascular dementia
I67.3	Binswanger's disease	Vascular dementia
F02.0	Dementia in Pick's disease	Rare dementia
F02.1	Dementia in Creutzfeldt-Jakob disease	Rare dementia
F02.2	Dementia in Huntington's disease	Rare dementia
F02.3	Dementia in Parkinson's disease	Rare dementia
F02.4	Dementia in human immunodef virus [HIV] disease	Rare dementia

ICD	LABEL	AGG
F02.8	Dementia in other specified diseases classified elsewhere	Rare dementia
F02	Dementia in other diseases classified elsewhere	Unspecified dementia type
F03	Unspecified dementia	Unspecified dementia type
F05.1	Delirium superimposed on dementia	Unspecified dementia type
F05.0	Delirium not superimposed on dementia, so described	Possible dementia
G31.0	Circumscribed brain atrophy	Possible dementia
G31.1	Senile degeneration of brain, not otherwise classified	Possible dementia
G31.2	Degeneration of nervous system due to alcohol	Possible dementia
G31.8	Other specified degenerative diseases of nervous system	Possible dementia
G31.9	Degenerative disease of nervous system, unspecified	Possible dementia
331.0	Alzheimer's disease	Alzheimer's disease
290.4	Vascular dementia	Vascular dementia
046.19	Creutzfeldt Jacob	Rare dementia
333.4	Huntingdon's	Rare dementia
331.1	Frontotemporal dementia	Rare dementia
290.0	Senile dementia, uncomplicated	Unspecified dementia type
290.1	Presenile dementia	Unspecified dementia type
290.2	Senile dementia with delusional features	Unspecified dementia type
290.3	Senile dementia with delirium	Unspecified dementia type
294.9	unspecified persistent mental disorders	Possible dementia
331.2	Senile degeneration	Possible dementia
331.9	Cerebral degeneration unspecified	Possible dementia

Table 15: Read codes used for dementia patient's identification in CPRD.

Read code	
	Alzheimer's disease
Eu00012	[X]Primary degen dementia, Alzheimer's type, presenile onset
Eu00013	[X]Alzheimer's disease type 2
Fyu3000	[X]Other Alzheimer's disease
Eu00111	[X]Alzheimer's disease type 1
Eu01111	[X]Predominantly cortical dementia
Eu00113	[X]Primary degen dementia of Alzheimer's type, senile onset
Eu00011	[X]Presenile dementia,Alzheimer's type
Eu00000	[X]Dementia in Alzheimer's disease with early onset
Eu00z00	[X]Dementia in Alzheimer's disease, unspecified
Eu00100	[X]Dementia in Alzheimer's disease with late onset
F110100	Alzheimer's disease with late onset
F110000	Alzheimer's disease with early onset
Eu00200	[X]Dementia in Alzheimer's dis, atypical or mixed type
Eu00112	[X]Senile dementia,Alzheimer's type
Eu00z11	[X]Alzheimer's dementia unspec
Eu00.00	[X]Dementia in Alzheimer's disease
F110.00	Alzheimer's disease
	Vascular dementia
E004100	Arteriosclerotic dementia with delirium
Eu01000	[X]Vascular dementia of acute onset

E004200	Arteriosclerotic dementia with paranoia
Eu01y00	[X]Other vascular dementia
E004300	Arteriosclerotic dementia with depression
Eu01200	[X]Subcortical vascular dementia
E004000	Uncomplicated arteriosclerotic dementia
Eu01300	[X]Mixed cortical and subcortical vascular dementia
Eu01z00	[X]Vascular dementia, unspecified
E004z00	Arteriosclerotic dementia NOS
Eu01.11	[X]Arteriosclerotic dementia
Eu01100	[X]Multi-infarct dementia
E004.00	Arteriosclerotic dementia
E004.11	Multi infarct dementia
Eu01.00	[X]Vascular dementia
	Rare dementia
Eu02400	[X]Dementia in human immunodef virus [HIV] disease
Eu02100	[X]Dementia in Creutzfeldt-Jakob disease
Eu02200	[X]Dementia in Huntington's disease
Eu02000	[X]Dementia in Pick's disease
F111.00	Pick's disease
Eu02300	[X]Dementia in Parkinson's disease
Eu02500	[X]Lewy body dementia
F116.00	Lewy body disease
	Unspecified dementia type
E002z00	Senile dementia with depressive or paranoid features NOS
Eu02z11	[X] Presenile dementia NOS
Eu02y00	[X]Dementia in other specified diseases classif elsewhere
E001000	Uncomplicated presenile dementia
E001100	Presenile dementia with delirium
E012.00	Other alcoholic dementia
Eu04100	[X]Delirium superimposed on dementia
Eu02z13	[X] Primary degenerative dementia NOS
E001300	Presenile dementia with depression
E001200	Presenile dementia with paranoia
Eu02z16	[X] Senile dementia, depressed or paranoid type

E001z00	Presenile dementia NOS
E002.00	Senile dementia with depressive or paranoid features
E003.00	Senile dementia with delirium
Eu10711	[X]Alcoholic dementia NOS
E012.11	Alcoholic dementia NOS
E041.00	Dementia in conditions EC
E002000	Senile dementia with paranoia
Eu02.00	[X]Dementia in other diseases classified elsewhere
E002100	Senile dementia with depression
E001.00	Presenile dementia
9hD0.00	Excepted from dementia quality indicators: Patient unsuitable
Eu02z14	[X] Senile dementia NOS
E000.00	Uncomplicated senile dementia
Eu02z00	[X] Unspecified dementia
E00..11	Senile dementia
E00..12	Senile/presenile dementia
E02y100	Drug-induced dementia
	Possible dementia
Eu04000	[X]Delirium not superimposed on dementia, so described

Table 16:: GitHub link to results of the selected studies and derived data variables in the literature review of dementia progression studies.

https://github.com/camcaan/LitReview/blob/master/LitReview_Table.xlsx

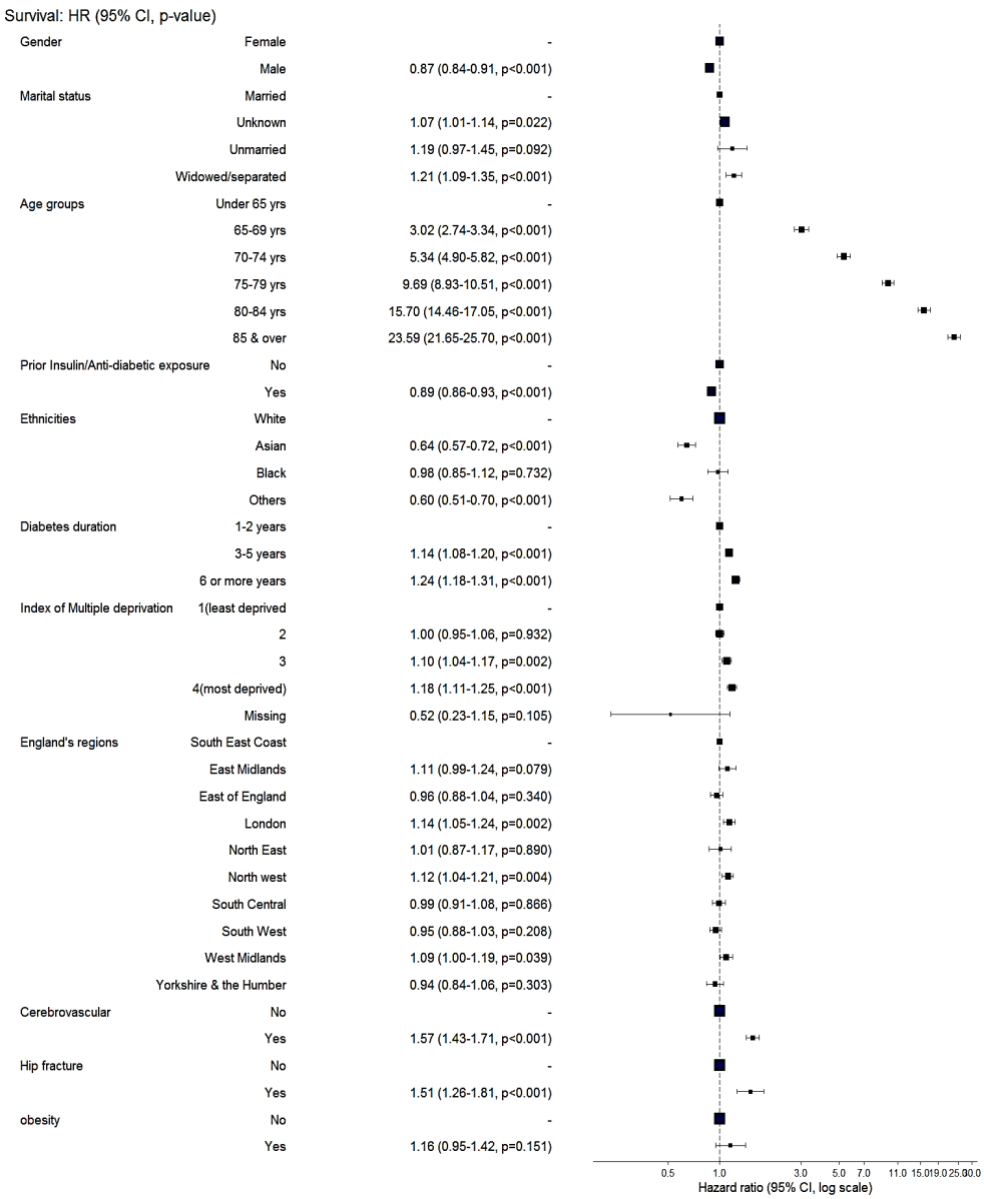


Figure 19: Hazard ratio plot for time to dementia in diabetic patients

Table 17 : Hazard Ratios (HRs) with 95% CI and P-values for each transition in the multi-state model

Term	estimate	std.error	robust.se	statistic	p.value	conf.low	conf.high
Dementia diagnosis --> Hospital (State 1 (s0) --> State 2)							
Home care	0.93	0.012	0.012	-6.04	1.56e-09	0.91	0.95
Age group 75-85 vs Under 75	1.34	0.014	0.014	20.95	1.74e-97	1.30	1.38
Age group 85+	1.90	0.016	0.016	40.52	0.00e+00	1.84	1.96
Male vs Female	1.12	0.011	0.011	10.67	1.34e-26	1.10	1.15
IMD = 2 vs 1 (least deprived)	1.04	0.013	0.013	2.96	3.06e-03	1.01	1.07
IMD = 3	1.08	0.014	0.014	5.73	9.97e-09	1.05	1.11
IMD = 4 (most deprived)	1.11	0.015	0.015	7.24	4.62e-13	1.08	1.15
Alzheimer's dementia vs Unspecified	0.86	0.010	0.010	-14.18	1.26e-45	0.85	0.88
Rare dementia	1.20	0.044	0.047	3.98	6.90e-05	1.10	1.32

Vascular dementia	1.06	0.023	0.024	2.48	1.32e-02	1.01	1.11
Prior Anti-cardiovascular drugs	1.17	0.010	0.011	14.99	9.02e-51	1.15	1.20
Prior anti-diabetic drugs	1.20	0.020	0.020	9.36	7.82e-21	1.16	1.25
Subsequent hospital admission from Out (State 3 -- > State 2)							
Home care	1.06	0.009	0.015	3.77	1.66e-04	1.03	1.09
Age group 75-85 vs Under 75	0.98	0.011	0.019	-1.20	2.32e-01	0.94	1.01
Age group 85+	0.96	0.013	0.021	-1.74	8.12e-02	0.92	1.00
Male vs Female	1.28	0.009	0.014	17.40	7.58e-68	1.25	1.32
IMD = 2 vs 1 (least deprived)	1.03	0.011	0.018	1.78	7.48e-02	1.00	1.07
IMD = 3	1.06	0.012	0.019	3.30	9.51e-04	1.03	1.10
IMD = 4 (most deprived)	1.17	0.012	0.020	8.17	3.02e-16	1.13	1.22

Alzheimer's dementia vs Unspecified	1.03	0.009	0.014	2.08	3.74e-02	1.00	1.06
Rare dementia	1.17	0.033	0.084	1.83	6.75e-02	0.99	1.38
Vascular dementia	1.03	0.017	0.027	1.21	2.28e-01	0.98	1.09
Prior Anti-cardiovascular drugs	0.95	0.010	0.015	-3.41	6.46e-04	0.92	0.98
Prior anti-diabetic drugs	1.11	0.015	0.027	3.90	9.57e-05	1.05	1.17
Mild frailty	1.22	0.014	0.023	8.74	2.24e-18	1.17	1.28
Moderate frailty	1.42	0.015	0.024	14.51	1.04e-47	1.36	1.49
Severe frailty	1.80	0.014	0.024	24.53	6.68e-133	1.72	1.89
30 day readmission	1.62	0.009	0.018	27.44	9.88e-166	1.56	1.68
Subsequent hospital admission from Institution (State 4 --> State 2)							

Home care	1.03	0.043	0.053	0.48	6.32e-01	0.92	1.14
Age group 75-85 vs Under 75	0.91	0.047	0.056	-1.76	7.86e-02	0.81	1.01
Age group 85+	0.85	0.057	0.068	-2.45	1.44e-02	0.74	0.97
Male vs Female	1.24	0.039	0.048	4.52	6.20e-06	1.13	1.36
IMD = 2 vs 1 (least deprived)	1.08	0.051	0.059	1.36	1.73e-01	0.97	1.21
IMD = 3	1.18	0.052	0.062	2.66	7.87e-03	1.04	1.33
IMD = 4 (most deprived)	1.35	0.051	0.062	4.82	1.46e-06	1.20	1.53
Alzheimer's dementia vs Unspecified	1.20	0.038	0.046	3.97	7.12e-05	1.10	1.31
Rare dementia	0.83	0.165	0.182	-1.05	2.96e-01	0.58	1.18
Vascular dementia	1.41	0.070	0.086	4.03	5.66e-05	1.19	1.68
Prior Anti-cardiovascular drugs	1.05	0.042	0.052	1.00	3.18e-01	0.95	1.17
Prior anti-diabetic drugs	1.29	0.069	0.088	2.89	3.88e-03	1.08	1.53

Mild frailty	1.43	0.059	0.066	5.43	5.54e-08	1.26	1.63
Moderate frailty	1.70	0.064	0.072	7.35	1.98e-13	1.47	1.96
Severe frailty	2.31	0.062	0.073	11.40	4.31e-30	2.00	2.66
30 day readmission	1.18	0.041	0.050	3.36	7.72e-04	1.07	1.31
Hospital discharge (State 2 --> State 3)							
Home care	1.29	0.008	0.028	9.18	4.38e-20	1.23	1.37
Age group 75-85 vs Under 75	0.99	0.010	0.035	-0.24	8.13e-01	0.93	1.06
Age group 85+	0.97	0.012	0.039	-0.74	4.62e-01	0.90	1.05
Male vs Female	1.03	0.008	0.026	1.11	2.66e-01	0.98	1.08
IMD = 2 vs 1 (least deprived)	1.12	0.010	0.034	3.23	1.24e-03	1.04	1.19
IMD = 3	1.11	0.010	0.035	2.83	4.61e-03	1.03	1.18
IMD = 4 (most deprived)	1.26	0.011	0.036	6.39	1.69e-10	1.17	1.35

Alzheimer's dementia vs Unspecified	1.20	0.008	0.026	7.08	1.44e-12	1.14	1.27
Rare dementia	0.77	0.030	0.111	-2.38	1.74e-02	0.62	0.95
Vascular dementia	1.07	0.016	0.047	1.43	1.54e-01	0.98	1.17
Prior Anti-cardiovascular drugs	1.06	0.009	0.030	1.86	6.24e-02	1.00	1.12
Prior anti-diabetic drugs	1.06	0.014	0.042	1.39	1.64e-01	0.98	1.15
Mild frailty	1.33	0.013	0.044	6.46	1.03e-10	1.22	1.45
Moderate frailty	1.69	0.013	0.045	11.55	7.46e-31	1.54	1.84
Severe frailty	2.05	0.013	0.044	16.24	2.65e-59	1.88	2.23
30 day readmission	1.09	0.008	0.029	2.87	4.14e-03	1.03	1.15
Institution to Out (State 4 -- > State 3)							
Home care	1.04	0.028	0.033	1.21	2.27e-01	0.98	1.11
Age group 75-85 vs Under 75	0.99	0.031	0.038	-0.39	6.95e-01	0.92	1.06

Age group 85+	0.94	0.037	0.044	-1.38	1.69e-01	0.86	1.03
Male vs Female	1.38	0.025	0.030	10.84	2.26e-27	1.30	1.46
IMD = 2 vs 1 (least deprived)	1.02	0.032	0.037	0.66	5.09e-01	0.95	1.10
IMD = 3	1.06	0.033	0.039	1.45	1.48e-01	0.98	1.14
IMD = 4 (most deprived)	1.12	0.033	0.040	2.79	5.20e-03	1.03	1.21
Alzheimer's dementia vs Unspecified	1.06	0.024	0.029	2.18	2.93e-02	1.01	1.13
Rare dementia	0.95	0.097	0.118	-0.48	6.33e-01	0.75	1.19
Vascular dementia	1.01	0.051	0.062	0.13	8.96e-01	0.89	1.14
Prior Anti-cardiovascular drugs	1.15	0.027	0.033	4.16	3.19e-05	1.07	1.22
Prior anti-diabetic drugs	1.22	0.047	0.059	3.39	7.12e-04	1.09	1.37
Mild frailty	1.29	0.036	0.040	6.25	4.13e-10	1.19	1.39
Moderate frailty	1.41	0.040	0.045	7.51	6.02e-14	1.29	1.54

Severe frailty	1.69	0.039	0.045	11.75	7.37e-32	1.55	1.85
30 day readmission	1.25	0.027	0.032	6.89	5.45e-12	1.17	1.33
Hospital to Institutionalisation (State 2 -- > State 4)							
Home care	1.07	0.019	0.032	2.13	3.29e-02	1.01	1.14
Age group 75-85 vs Under 75	1.15	0.022	0.038	3.72	2.00e-04	1.07	1.24
Age group 85+	1.23	0.025	0.043	4.84	1.30e-06	1.13	1.34
Male vs Female	0.94	0.017	0.029	-1.99	4.70e-02	0.89	1.00
IMD = 2 vs 1 (least deprived)	1.14	0.021	0.038	3.58	3.50e-04	1.06	1.23
IMD = 3	1.09	0.022	0.039	2.21	2.72e-02	1.01	1.18
IMD = 4 (most deprived)	1.30	0.023	0.041	6.36	1.97e-10	1.20	1.40
Alzheimer's dementia vs Unspecified	1.12	0.017	0.029	3.81	1.39e-04	1.06	1.18

Rare dementia	0.77	0.066	0.108	-2.42	1.56e-02	0.62	0.95
Vascular dementia	1.03	0.035	0.055	0.61	5.40e-01	0.93	1.15
Prior Anti-cardiovascular drugs	0.89	0.019	0.032	-3.49	4.84e-04	0.84	0.95
Prior anti-diabetic drugs	0.98	0.033	0.051	-0.31	7.59e-01	0.89	1.09
Mild frailty	1.21	0.025	0.046	4.17	3.11e-05	1.11	1.33
Moderate frailty	1.51	0.027	0.048	8.57	1.01e-17	1.37	1.66
Severe frailty	1.65	0.027	0.047	10.73	7.01e-27	1.51	1.81
30 day readmission	0.95	0.019	0.033	-1.70	8.88e-02	0.89	1.01
Admission due to injury	1.71	0.040	0.083	6.44	1.22e-10	1.45	2.01
Long hospital stay (>= 12 days)	1.34	0.026	0.049	5.98	2.20e-09	1.22	1.48
Out to Institution (State 3 --> State 4)							
Home care	0.76	0.062	0.077	-3.64	2.76e-04	0.65	0.88

Age group 75-85 vs Under 75	0.91	0.065	0.077	-1.16	2.44e-01	0.79	1.06
Age group 85+	0.79	0.079	0.090	-2.58	9.84e-03	0.66	0.95
Male vs Female	0.92	0.055	0.068	-1.19	2.35e-01	0.81	1.05
IMD = 2 vs 1 (least deprived)	1.16	0.072	0.082	1.77	7.72e-02	0.98	1.36
IMD = 3	1.16	0.075	0.087	1.76	7.91e-02	0.98	1.38
IMD = 4 (most deprived)	1.71	0.072	0.088	6.11	9.71e-10	1.44	2.04
Alzheimer's dementia vs Unspecified	1.28	0.053	0.064	3.83	1.27e-04	1.13	1.45
Rare dementia	0.92	0.240	0.257	-0.31	7.60e-01	0.56	1.53
Vascular dementia	1.68	0.094	0.119	4.37	1.22e-05	1.33	2.13
Prior Anti-cardiovascular drugs	0.65	0.059	0.071	-6.08	1.22e-09	0.56	0.75
Prior anti-diabetic drugs	1.30	0.093	0.116	2.27	2.35e-02	1.04	1.64
Mild frailty	1.24	0.082	0.089	2.43	1.51e-02	1.04	1.48

Moderate frailty	1.50	0.088	0.095	4.31	1.61e-05	1.25	1.81
Severe frailty	1.85	0.085	0.097	6.32	2.58e-10	1.53	2.24
30 day readmission	1.82	0.057	0.069	8.67	4.22e-18	1.59	2.08
Death inside hospital (State 2 -- > State 5)							
Home care	0.75	0.025	0.036	-7.85	4.09e-15	0.70	0.81
Age group 75-85 vs Under 75	1.48	0.029	0.045	8.67	4.44e-18	1.35	1.61
Age group 85+	2.16	0.032	0.050	15.37	2.63e-53	1.96	2.38
Male vs Female	1.38	0.020	0.032	10.12	4.40e-24	1.29	1.47
IMD = 2 vs 1 (least deprived)	1.10	0.026	0.042	2.27	2.30e-02	1.01	1.19
IMD = 3	1.10	0.027	0.043	2.34	1.95e-02	1.02	1.20
IMD = 4 (most deprived)	1.23	0.028	0.045	4.50	6.74e-06	1.12	1.34

Alzheimer's dementia vs Unspecified	0.94	0.020	0.033	-1.98	4.80e-02	0.88	1.00
Rare dementia	0.69	0.081	0.120	-3.13	1.73e-03	0.54	0.87
Vascular dementia	0.94	0.041	0.059	-1.06	2.88e-01	0.84	1.05
Prior Anti-cardiovascular drugs	1.28	0.023	0.037	6.72	1.83e-11	1.19	1.38
Prior anti-diabetic drugs	0.92	0.037	0.055	-1.47	1.42e-01	0.83	1.03
Mild frailty	1.99	0.041	0.054	12.77	2.31e-37	1.79	2.21
Moderate frailty	2.75	0.042	0.057	17.65	9.43e-70	2.46	3.08
Severe frailty	3.54	0.041	0.055	22.91	4.06e-116	3.18	3.95
30 day readmission	1.04	0.022	0.036	1.21	2.26e-01	0.97	1.12
Admission due to injury	1.30	0.053	0.093	2.86	4.28e-03	1.09	1.56
Long hospital stay (>= 12 days)	1.20	0.032	0.055	3.37	7.42e-04	1.08	1.34

Death outside hospital (State 3 --> State 5)							
Home care	0.68	0.019	0.020	-19.20	3.72e-82	0.66	0.71
Age group 75-85 vs Under 75	1.54	0.023	0.024	17.66	8.94e-70	1.47	1.62
Age group 85+	2.46	0.026	0.028	32.47	3.01e-231	2.33	2.60
Male vs Female	1.23	0.017	0.019	10.94	7.31e-28	1.18	1.28
IMD = 2 vs 1 (least deprived)	1.02	0.020	0.022	0.85	3.97e-01	0.98	1.06
IMD = 3	0.96	0.022	0.023	-1.96	5.03e-02	0.91	1.00
IMD = 4 (most deprived)	0.96	0.023	0.025	-1.82	6.81e-02	0.91	1.00
Alzheimer's dementia vs Unspecified	0.96	0.016	0.017	-2.50	1.23e-02	0.93	0.99
Rare dementia	1.15	0.070	0.071	2.00	4.51e-02	1.00	1.32

Vascular dementia	0.92	0.035	0.040	-2.01	4.48e-02	0.85	1.00
Prior Anti-cardiovascular drugs	0.97	0.018	0.019	-1.43	1.54e-01	0.94	1.01
Prior anti-diabetic drugs	1.01	0.033	0.036	0.40	6.87e-01	0.95	1.09
Mild frailty	1.42	0.028	0.027	12.79	1.89e-37	1.35	1.50
Moderate frailty	1.81	0.029	0.029	20.32	7.77e-92	1.71	1.92
Severe frailty	2.13	0.029	0.029	26.51	7.49e-155	2.01	2.25
30 day readmission	1.32	0.018	0.020	13.91	5.89e-44	1.27	1.38
Death inside Institution (State 4 --> State 5)							
Home care	0.87	0.032	0.034	-4.19	2.78e-05	0.81	0.93
Age group 75-85 vs Under 75	1.33	0.038	0.041	7.04	1.91e-12	1.23	1.44
Age group 85+	1.96	0.043	0.046	14.64	1.46e-48	1.79	2.14

Male vs Female	1.47	0.028	0.031	12.66	1.02e-36	1.39	1.57
IMD = 2 vs 1 (least deprived)	0.96	0.034	0.036	-1.24	2.15e-01	0.89	1.03
IMD = 3	0.89	0.036	0.039	-3.13	1.78e-03	0.82	0.96
IMD = 4 (most deprived)	0.84	0.038	0.040	-4.47	7.75e-06	0.77	0.90
Alzheimer's dementia vs Unspecified	0.94	0.027	0.029	-1.99	4.61e-02	0.89	1.00
Rare dementia	1.09	0.107	0.110	0.77	4.39e-01	0.88	1.35
Vascular dementia	0.89	0.058	0.063	-1.85	6.47e-02	0.79	1.01
Prior Anti-cardiovascular drugs	1.22	0.030	0.033	6.04	1.56e-09	1.14	1.30
Prior anti-diabetic drugs	0.90	0.060	0.069	-1.56	1.18e-01	0.79	1.03
Mild frailty	1.38	0.043	0.041	7.85	4.30e-15	1.28	1.50
Moderate frailty	1.85	0.045	0.045	13.62	2.97e-42	1.69	2.02
Severe frailty	2.06	0.044	0.046	15.80	3.18e-56	1.88	2.25

30 day readmission	1.09	0.030	0.033	2.51	1.21e-02	1.02	1.16

Supplements

For this review we limited ourselves to the evidence available from OVID MEDLINE for 2008 till April 2018 using the following search strategy (Search conducted on April 09, 2019):

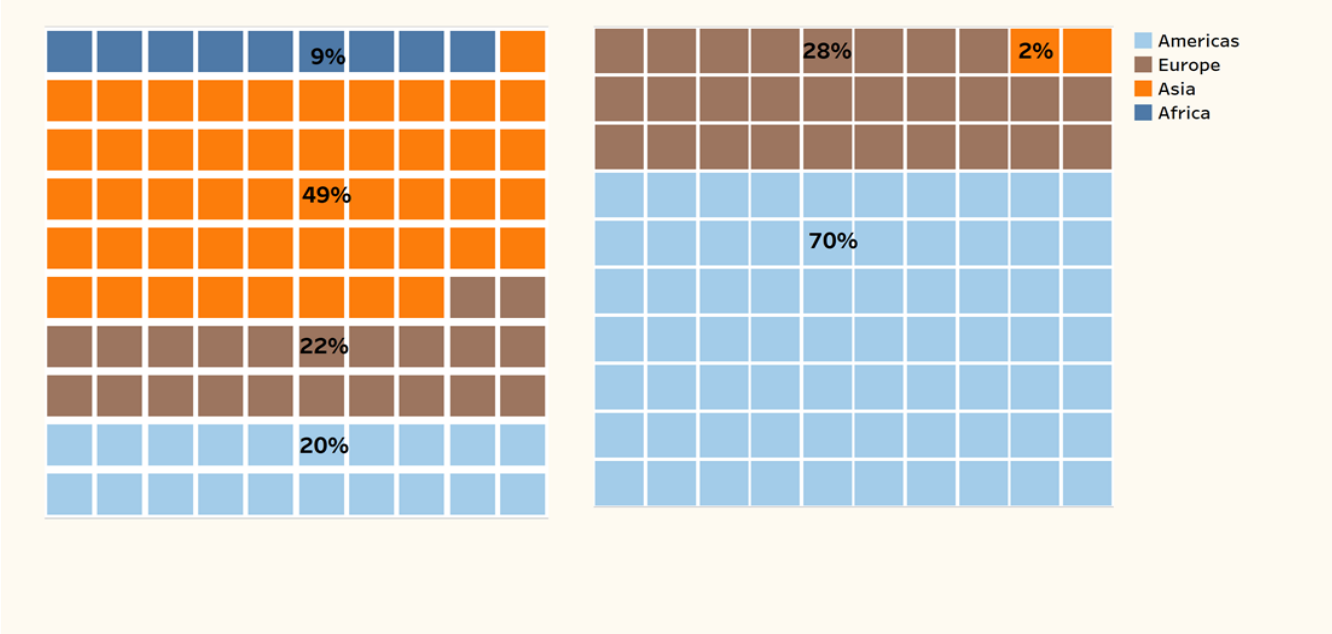
Search strategy for literature review

1 exp DEMENTIA/ (152924)
2 exp ALZHEIMER DISEASE/ (86571)
3 dementia.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (114720)
4 alzheim*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (145765)
5 *dementia/ (36898)
6 dementia.mp. (114720)
7 (alzheimer adj5 disease).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (92632)
8 (alzheimer adj5 amyloid).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (1578)
9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (234043)
10 exp Disease Progression/ (163579)
11 disease progression.mp. (197067)
12 (disease adj3 progression).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (220094)
13 exp disease course/ or disease course.mp. (11935)
14 (disease adj3 course).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (42497)
15 exp prognosis/ or prognosis.mp. or prognos*.mp. or progres*.mp. (2657125)
16 clinical course.mp. (57744)
17 10 or 11 or 12 or 13 or 14 or 15 or 16 (2725449)

18 9 and 17 (45249)
19 exp prospective studies/ (498887)
20 (prospective adj5 cohort).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (67993)
21 prospective studies.mp. (526388)
22 exp longitudinal studies/ or longitudinal.mp. (273221)
23 "observational studies".ti. or epidemiological studies/ (10658)
24 (cohort adj5 (study or studies or analy$)).mp. (383449)
25 ((follow-up adj5 (study or studies)) or longitudinal or retrospective or (observational adj5 (study or studies))).af. (1746809)
26 19 or 20 or 21 or 22 or 23 or 24 or 25 (2276908)
27 18 and 26 (8662)
28 (health utili#ation or healthcare utili#ation or health system utili#ation).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (4449)
29 27 or 28 (13105)
30 limit 29 to (english language and humans and "all aged (65 and over)" and last 10 years) (4824)
31 (201804* or 201805* or 201806* or 201807* or 201808* or 201809* or 201810* or 201811* or 201812* or 2019*).dt,ez,ed. (2018287)
32 30 and 31 (677) This is the result from the updated search since the last time the search was done in 31/03/2018
The final number of journals therefore was 4824+677= 5,501

****

E-figures : **Supplementary figures associated with literature review.**



(A) (B)

Fig (e1): Study locations and sample size percentage

*Asia also include Australia in chart B.

A) Percentage of global prevalence of dementia in different regions of the world (World Alzheimer's report, 2015) https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf.
This chart shows us how much in terms of percentage each region of the world contributes to global prevalence of Alzheimer's dementia.

B) Geographical distribution of dementia total sample size (N) in the studies identified in our review.
The waffle chart shows that 70% of the dementia cohort in the selected studies in this review were from Americas, while 28% were from Europe and Scandinavia and only 2% from Asia*(Also include Australia here).

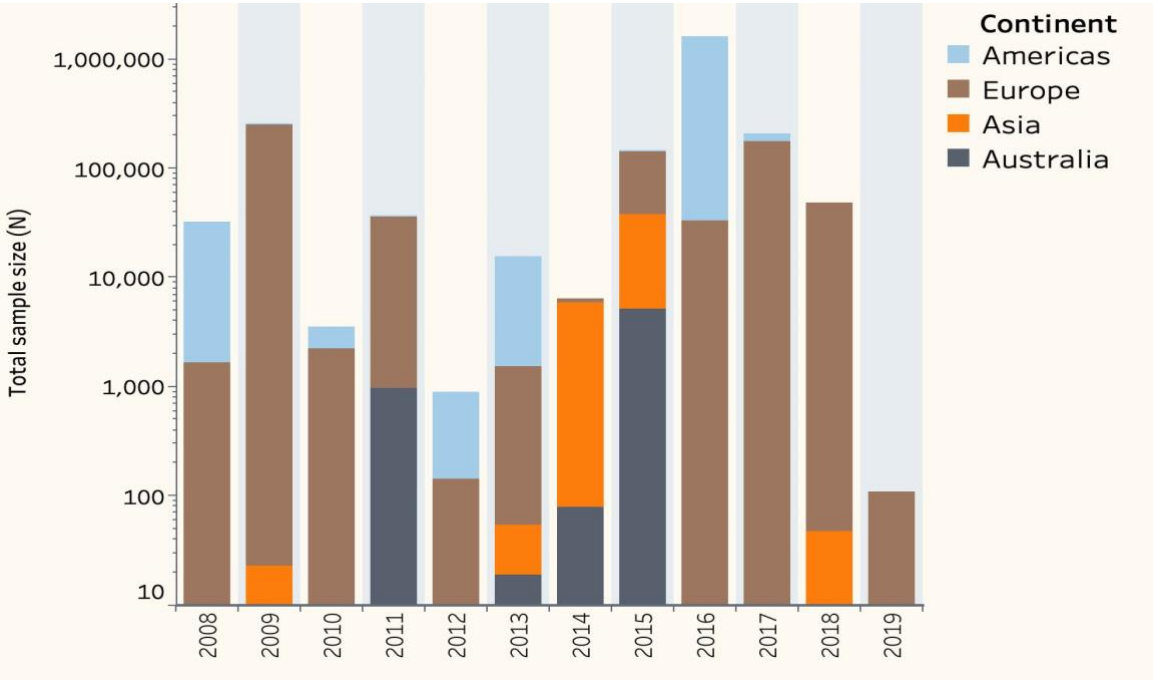


Fig (e2): Global distribution of studies in our review and total sample size in each year of publication.
The length of the bars shows the sum of sample size of all the studies which fulfilled our review inclusion criteria, published each year from different regions of the World.
Majority of the studies were originated from Americas and Europe.
Only one study with the largest sample size was selected when one study leads to multiple publications.

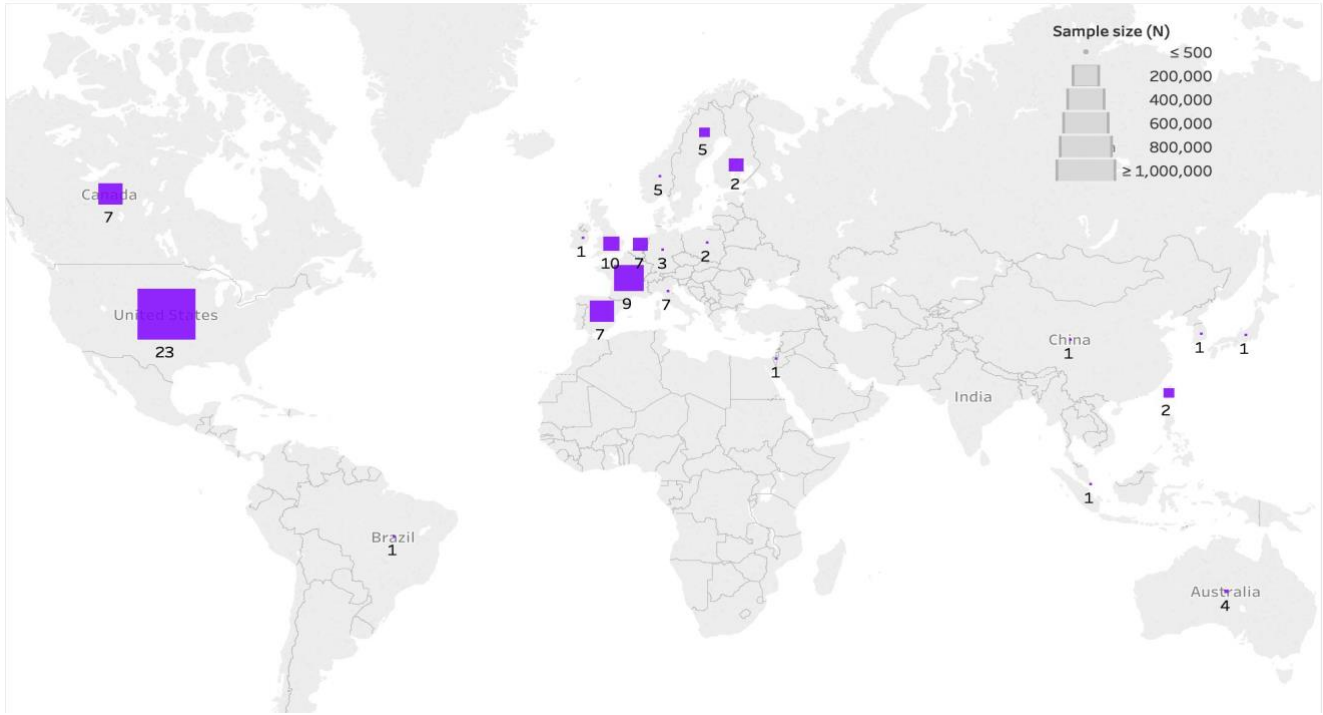
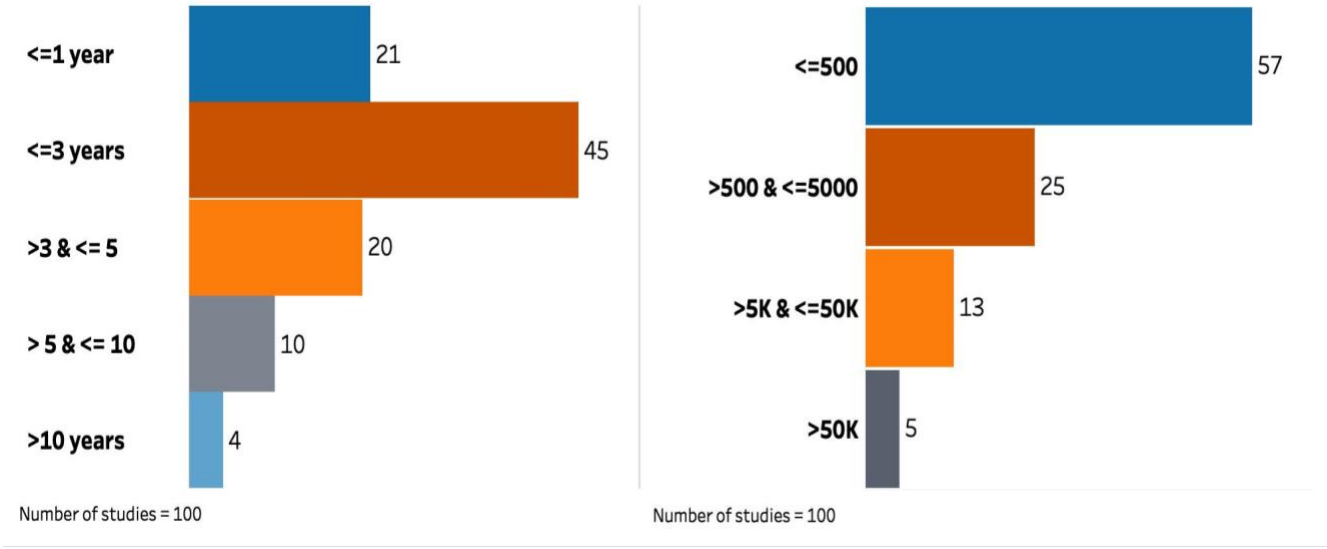


Figure (e3): Global distribution of studies identified in our systematic review. (total studies = 100)

The size of the squares represents the sample size (N) from each country while the numbers represent the number of studies. The map shows that majority of the studies were from North America and Europe.



(A) Average follow-up duration (Number of studies) (B) Study sample size (Number of studies)

Figure e4: (A) Number of studies with the average follow-up
(B) Number of studies and their sample size (N)

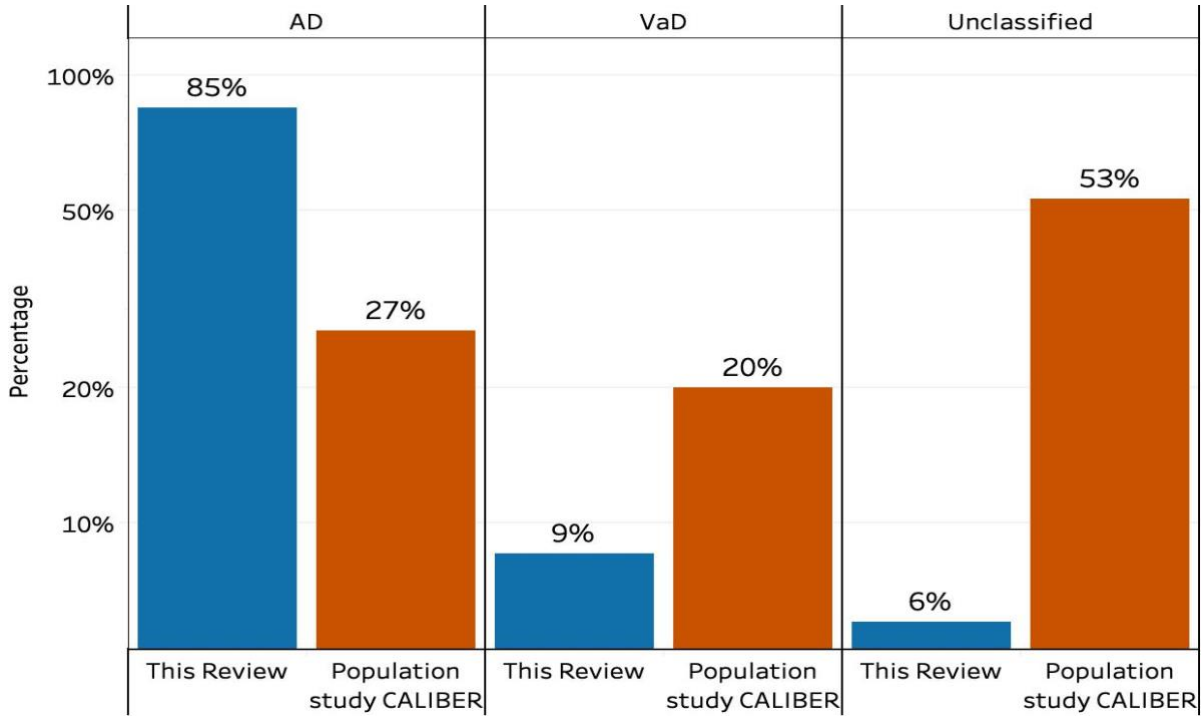


Figure (e5): Dementia subtypes reported in the studies compared to the English EHRs using CALIBER platform.

* CALIBER (CArdiovascular disease research using LInked Bespoke studies and Electronic health Records) is a platform combining the English primary care data (CPRD) with mortality data and hospital data.
The bar graph shows the difference in prevalence of dementia subtypes reported in the population study using CALIBER platform and studies in this review. In this review it clearly shows that Vascular dementia (VaD) is underrepresented, and cases were mainly identified as Alzheimer's dementia (AD). The data is from a study by Mar Pujadas-Rodriguez, et al, 2017 who used CALIBER platform to identify dementia prognosis. Total 47,386 individuals with dementia were identified.

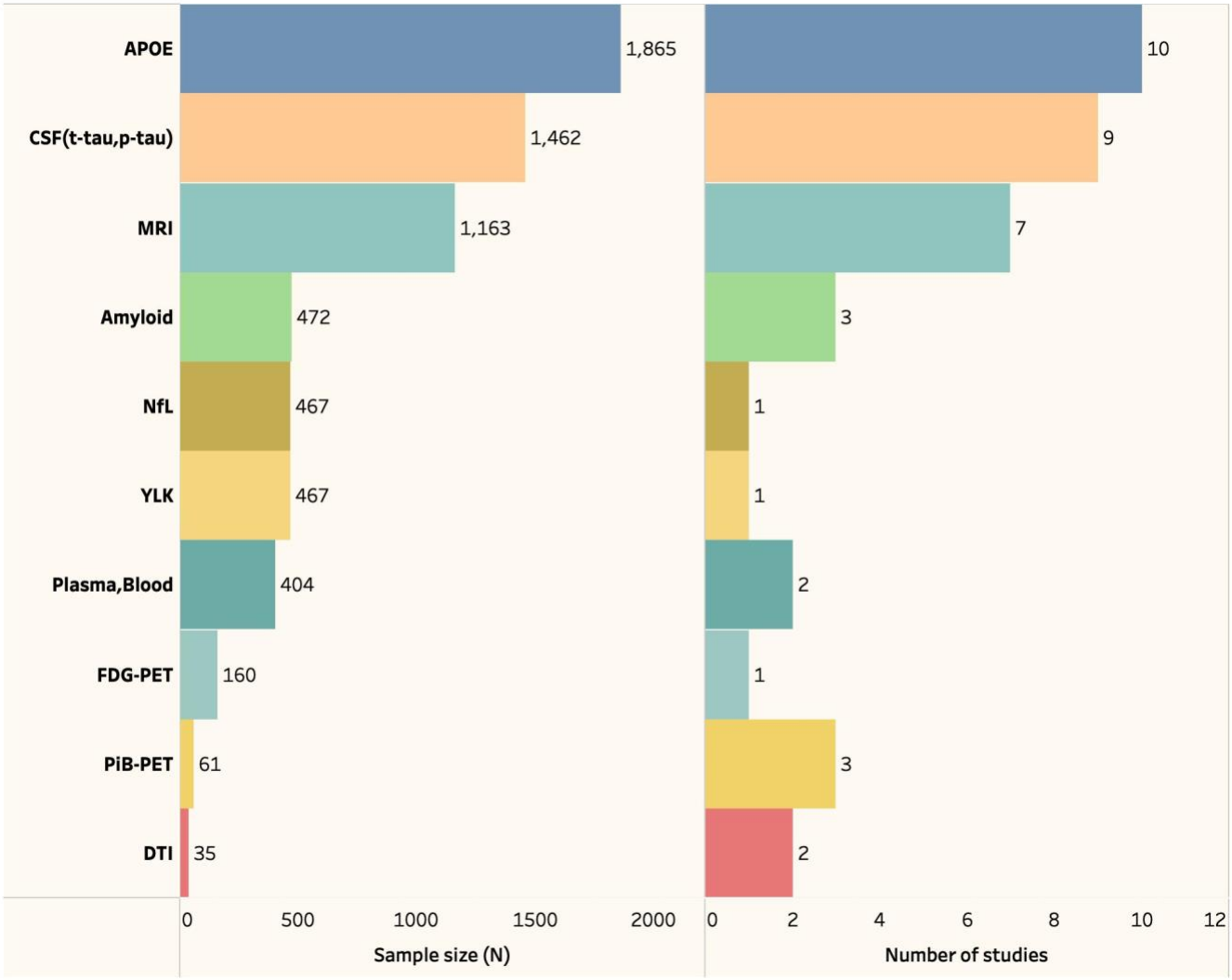


Fig (e6): Number of studies measuring different biomarkers with their sample size.

The length of bars on the right represents the number of studies reporting each biomarker and the bars on the left represents the total sample size in those studies.

APOE: Apolipo protein E
PiB: Pittsburgh compound B
DTI: Diffusion tensor imaging
MRI: Magnetic resonance imaging
FDG-PET: Fluoro-deoxy-D-glucose- Positron Emission Tomography
NfL: Neurofilament light
YLK-40: name given to an astrocytic inflammatory protein.

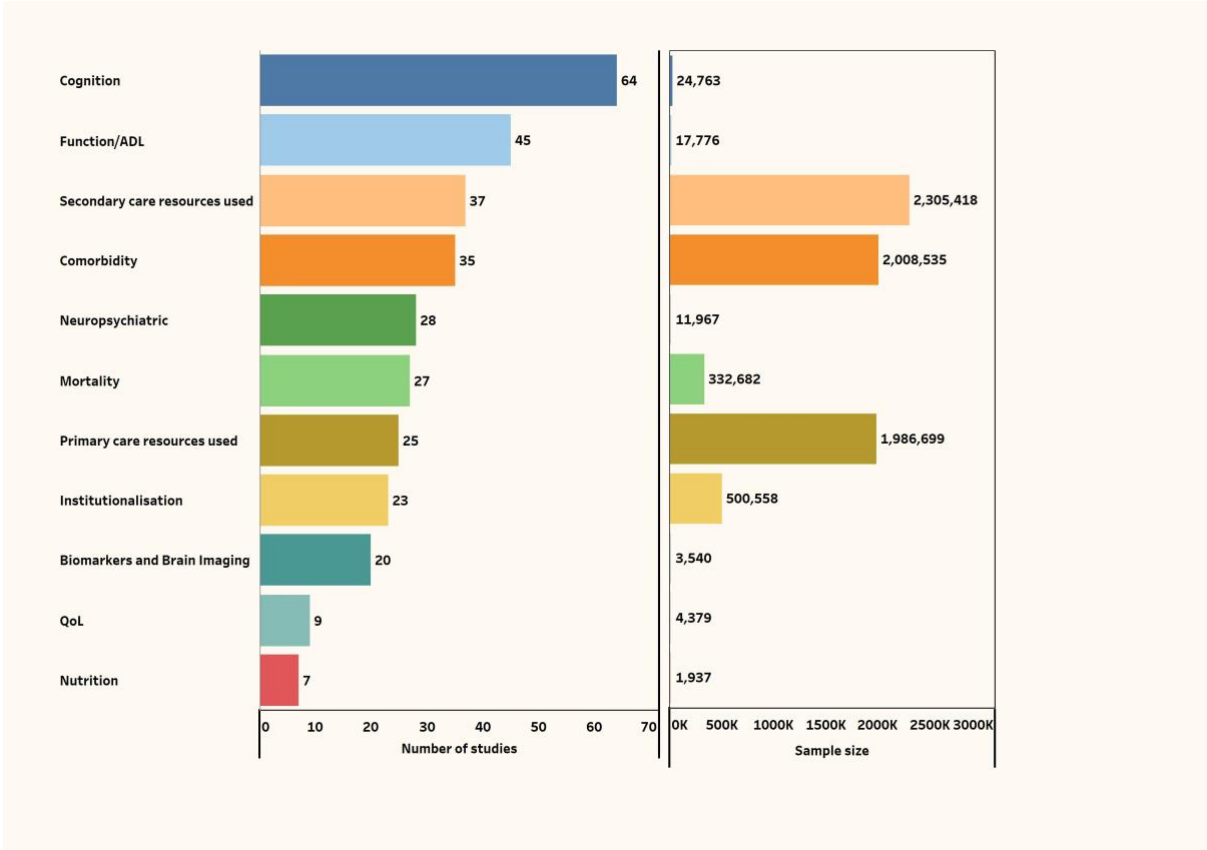


Figure (e7): Outcome measures with the number of studies reporting it and their sample size (N)
Note: One study can report more than one of the outcomes in the chart.
This bar chart shows the number of studies reporting each of the outcome measure and the associated sample size of the individuals followed in those studies.
We can see that cognition was reportedly measured in 64 studies, however, these studies had small sample size and the total N was only ~24k in these studies. On the other hand, there were less studies measuring healthcare resource utilisation, but they had large N because of their use of health insurance databases.

Supplementary figures associated with chapter 8 multi-state model

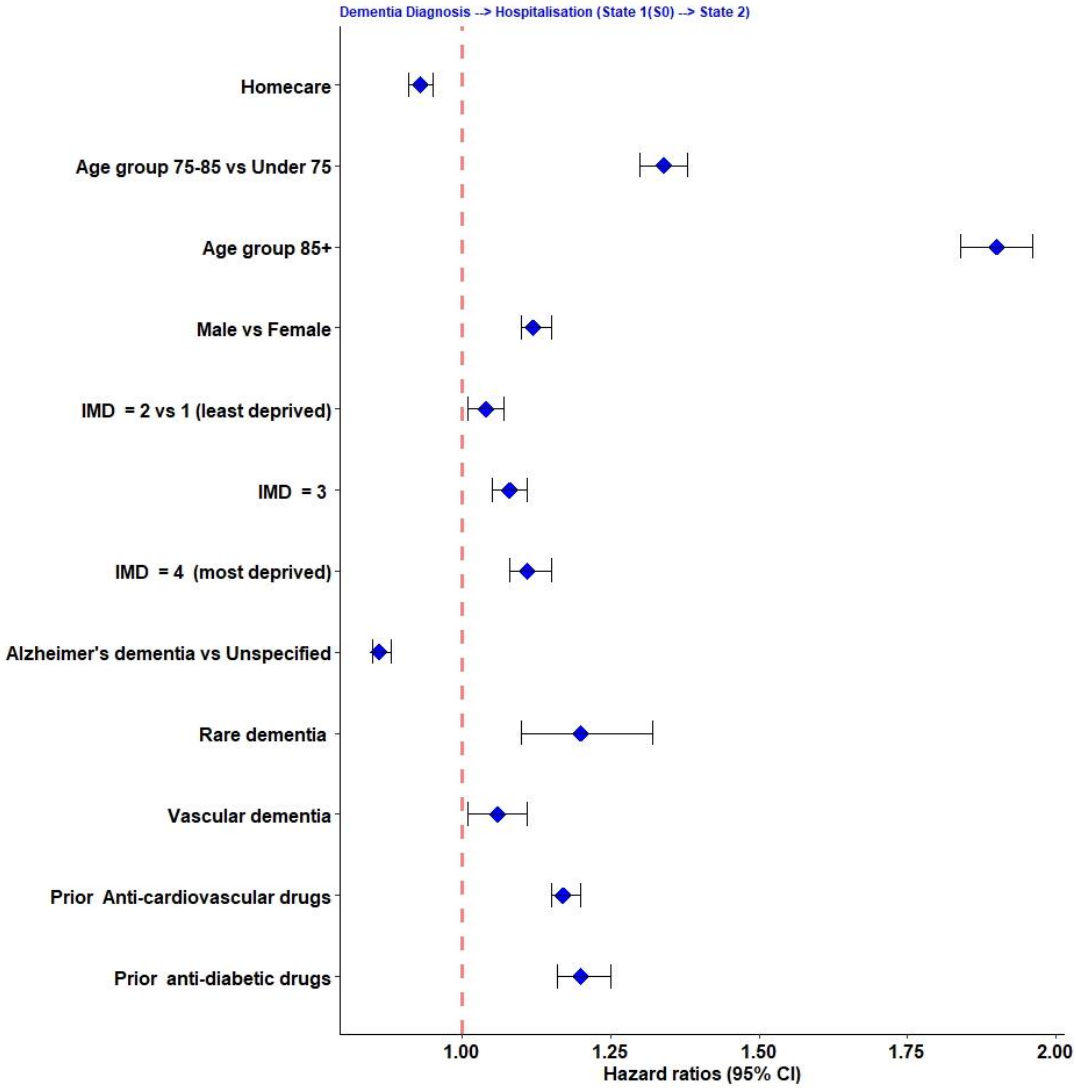


Figure e8: Hazard plot showing the hazard ratios of hospitalisation after dementia diagnosis.

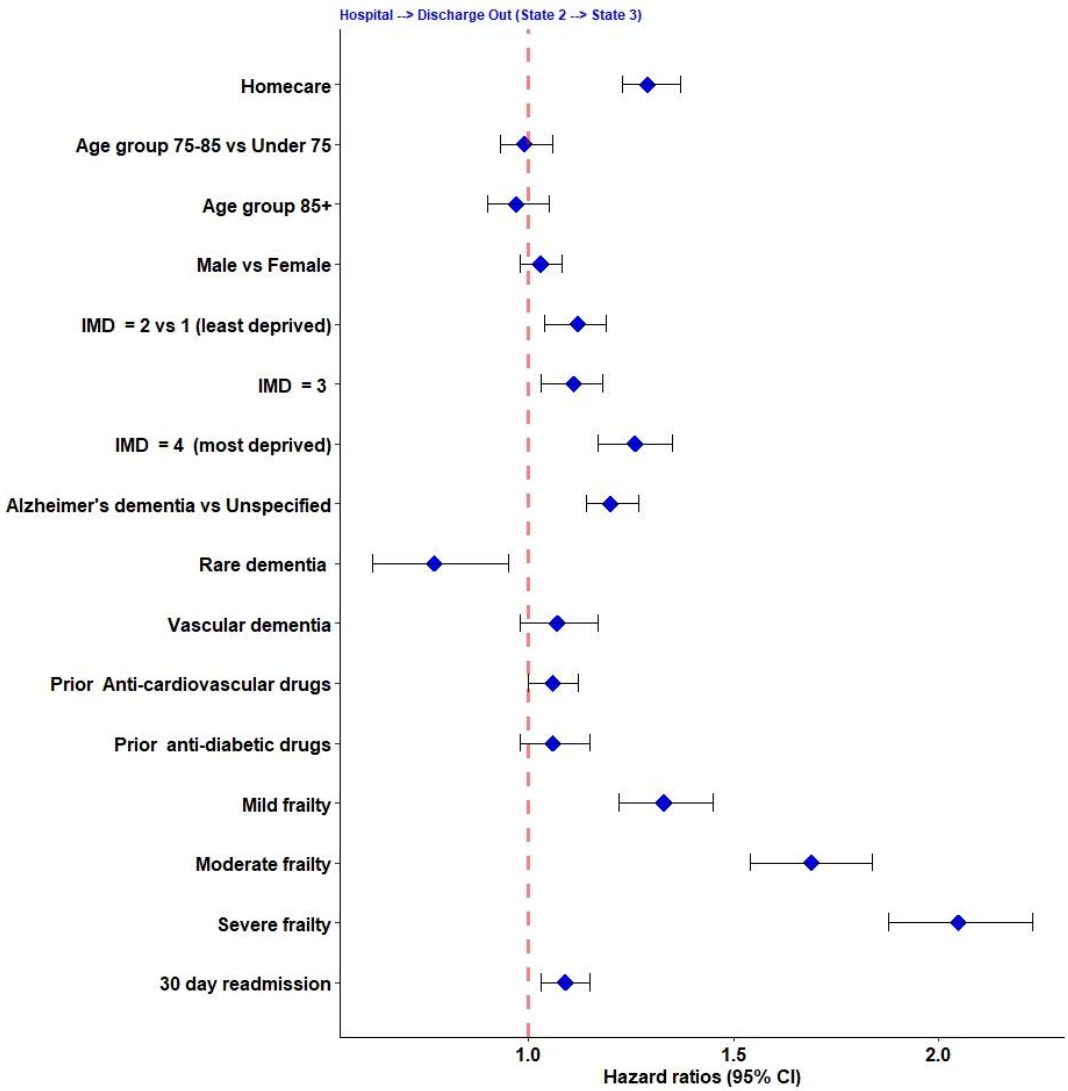


Figure e9: Hazard plot showing the hazard ratios of hospital discharge to usual place of living

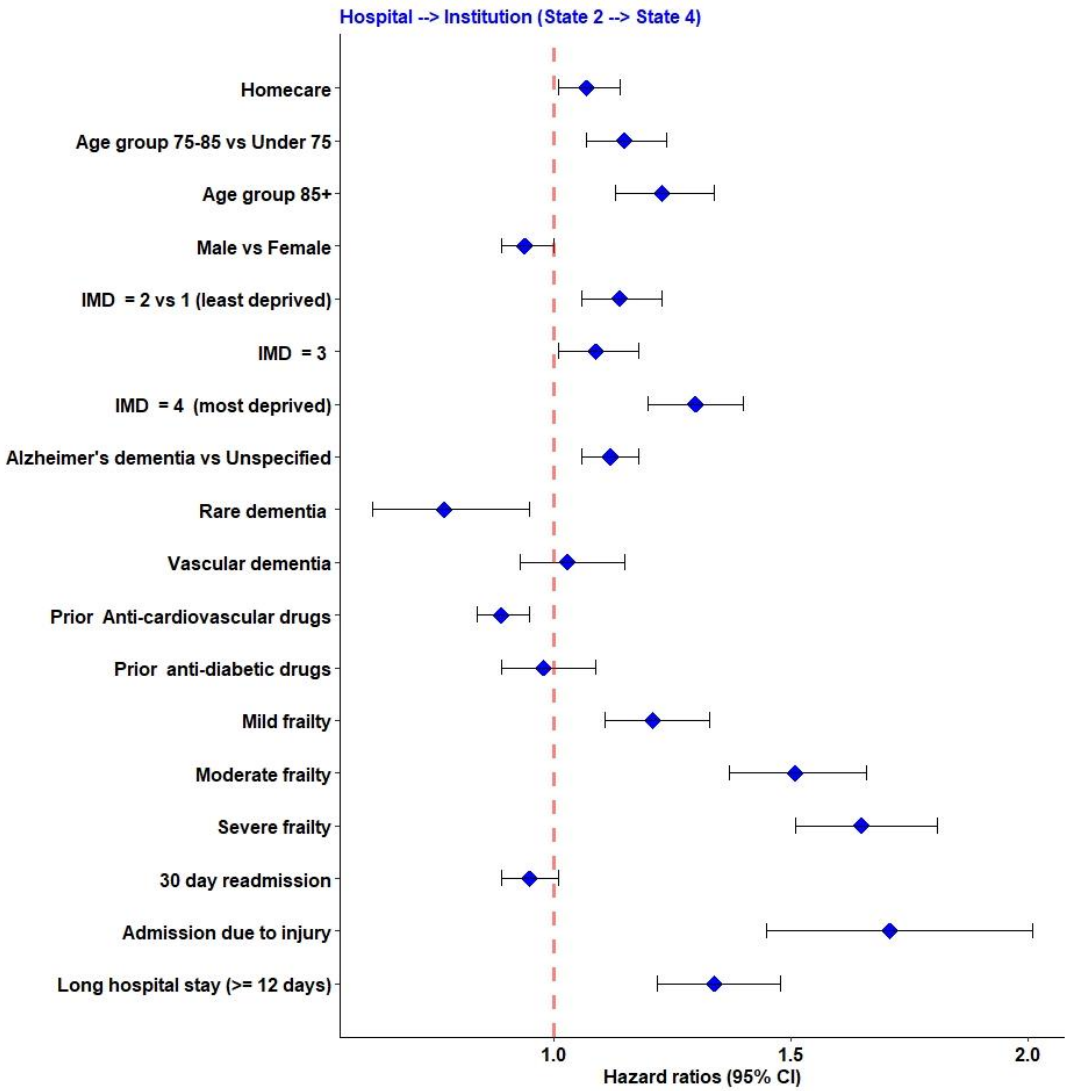


Figure e10: Hazard plot showing the hazard ratios of hospital discharge to institutional care

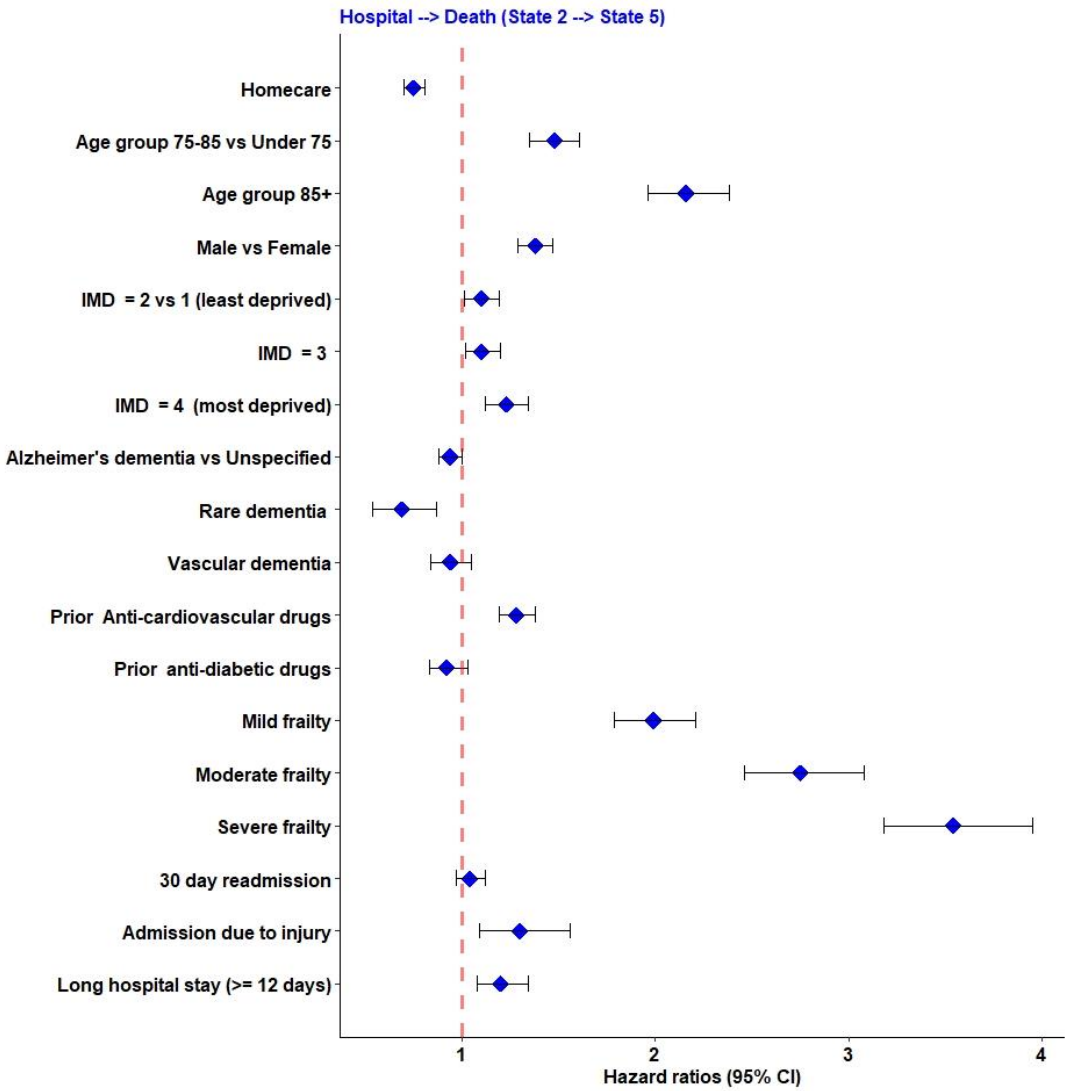


Figure e11: Hazard plot showing the hazard ratios of in-hospital deaths

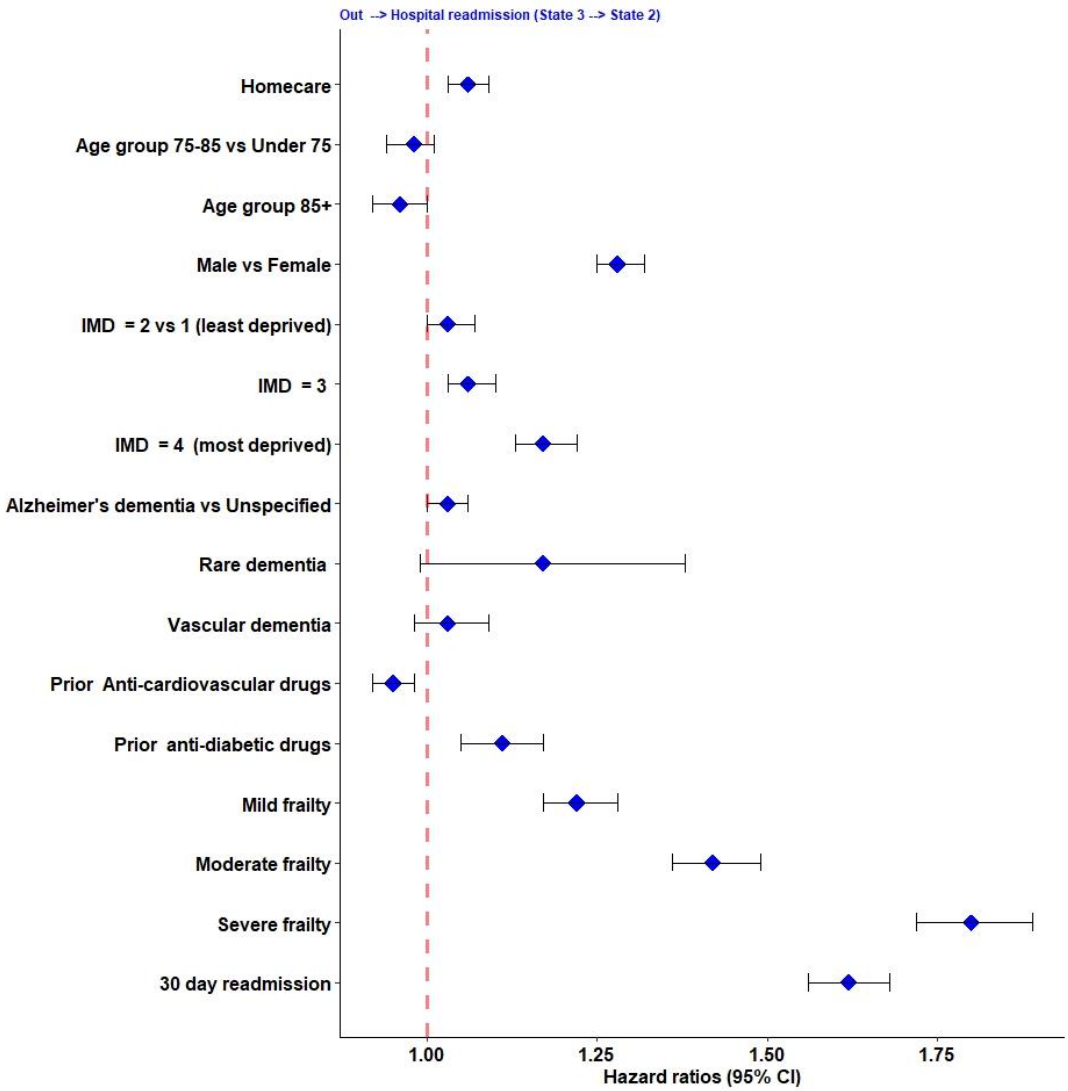


Figure e12: Hazard plot showing the hazard ratios of rehospitalisation once discharged home

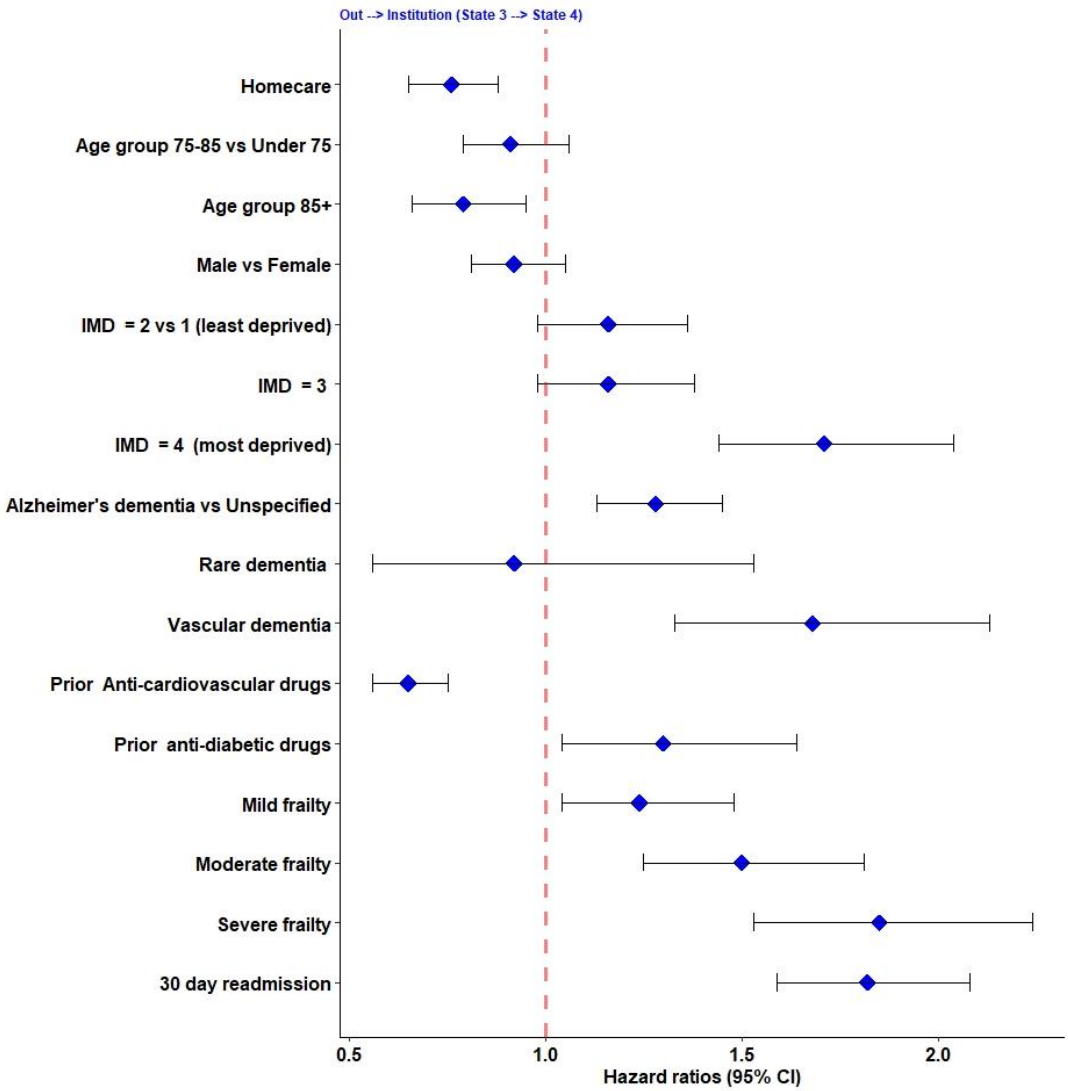


Figure e13: Hazard plot showing the hazard ratios of institutionalisation from home

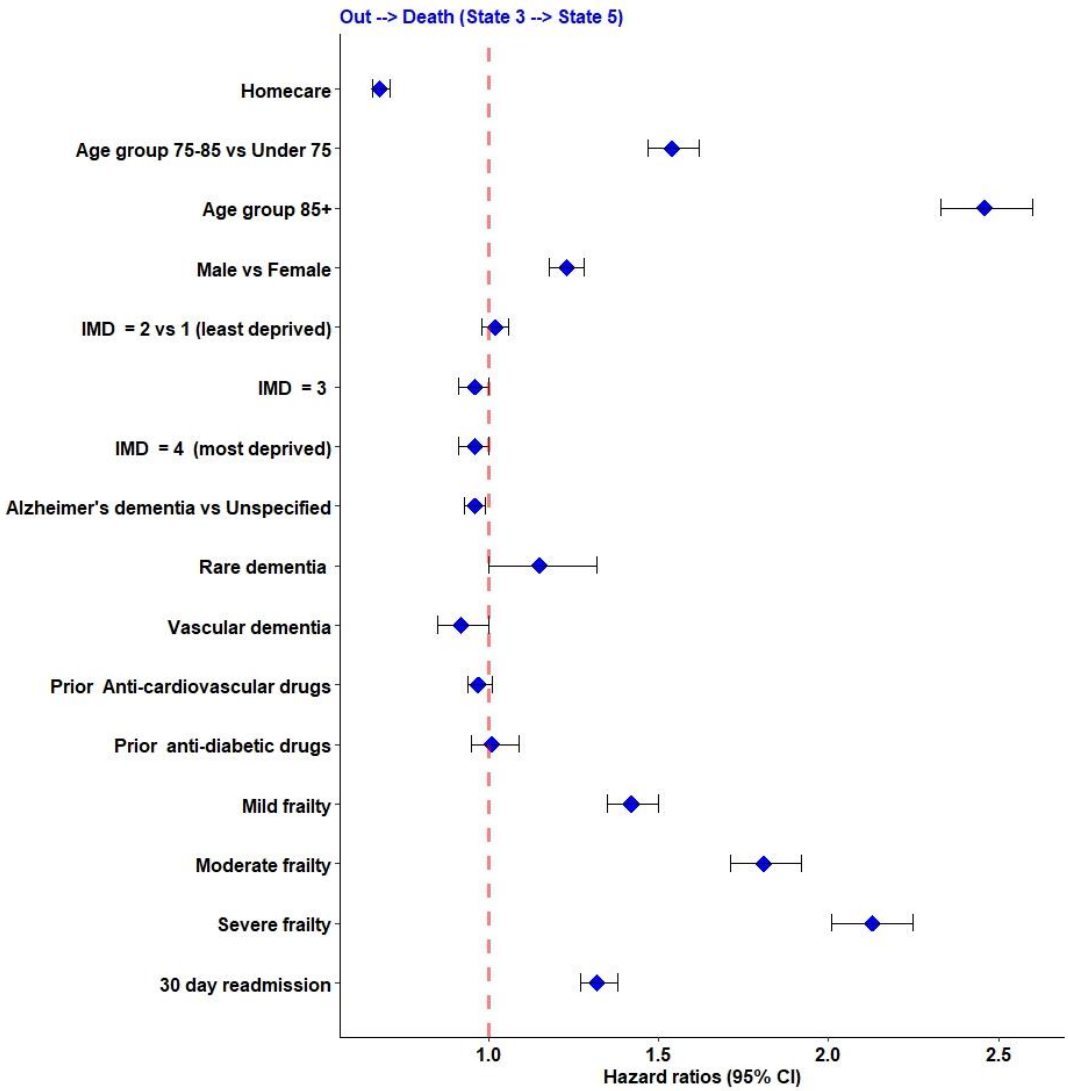


Figure e14: Hazard plot showing the hazard ratios of death outside hospital

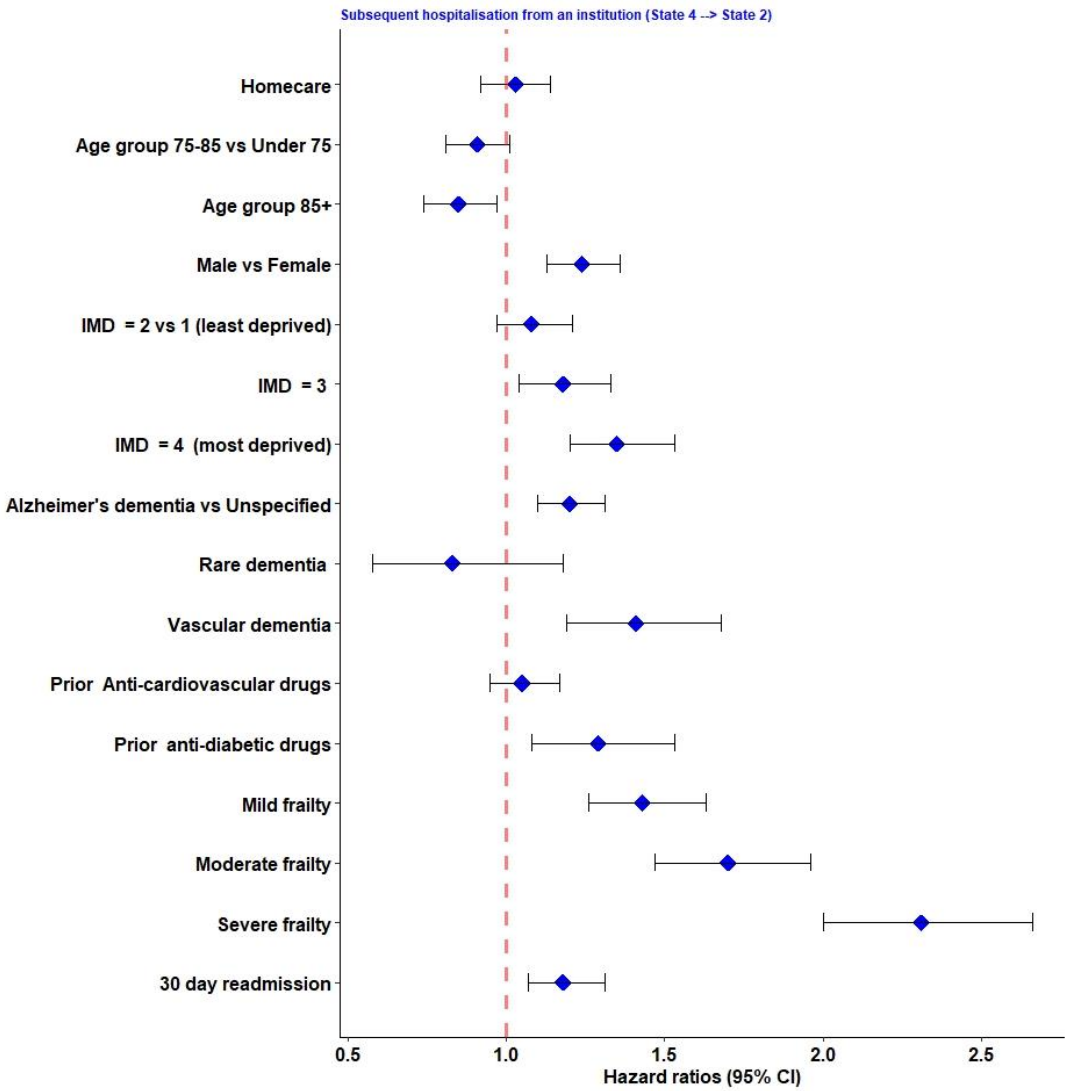


Figure e15: Hazard plot showing the hazard ratios of rehospitalisation from long term care institutions

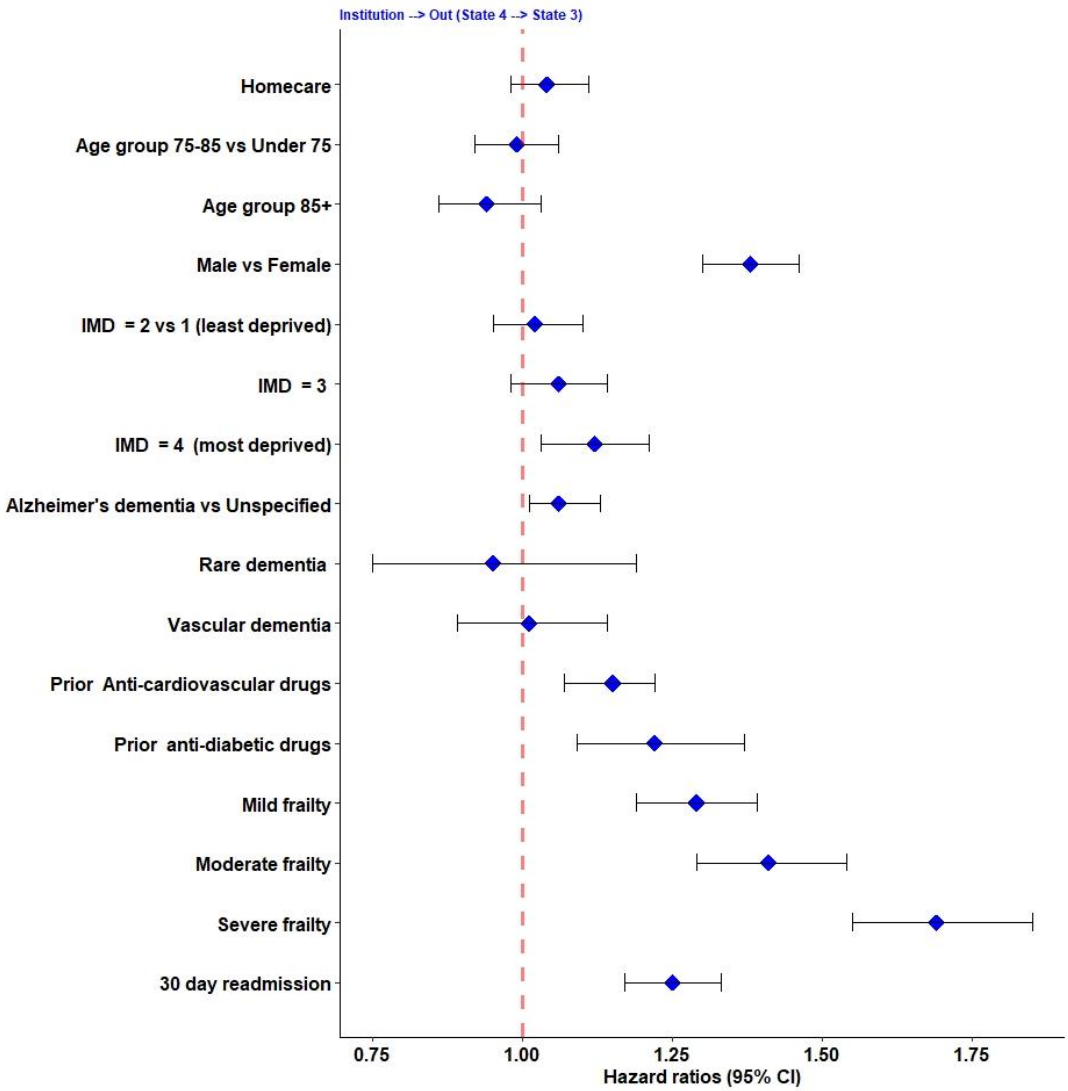


Figure e16: Hazard plot showing the hazard ratios discharge home from long-term care institutions

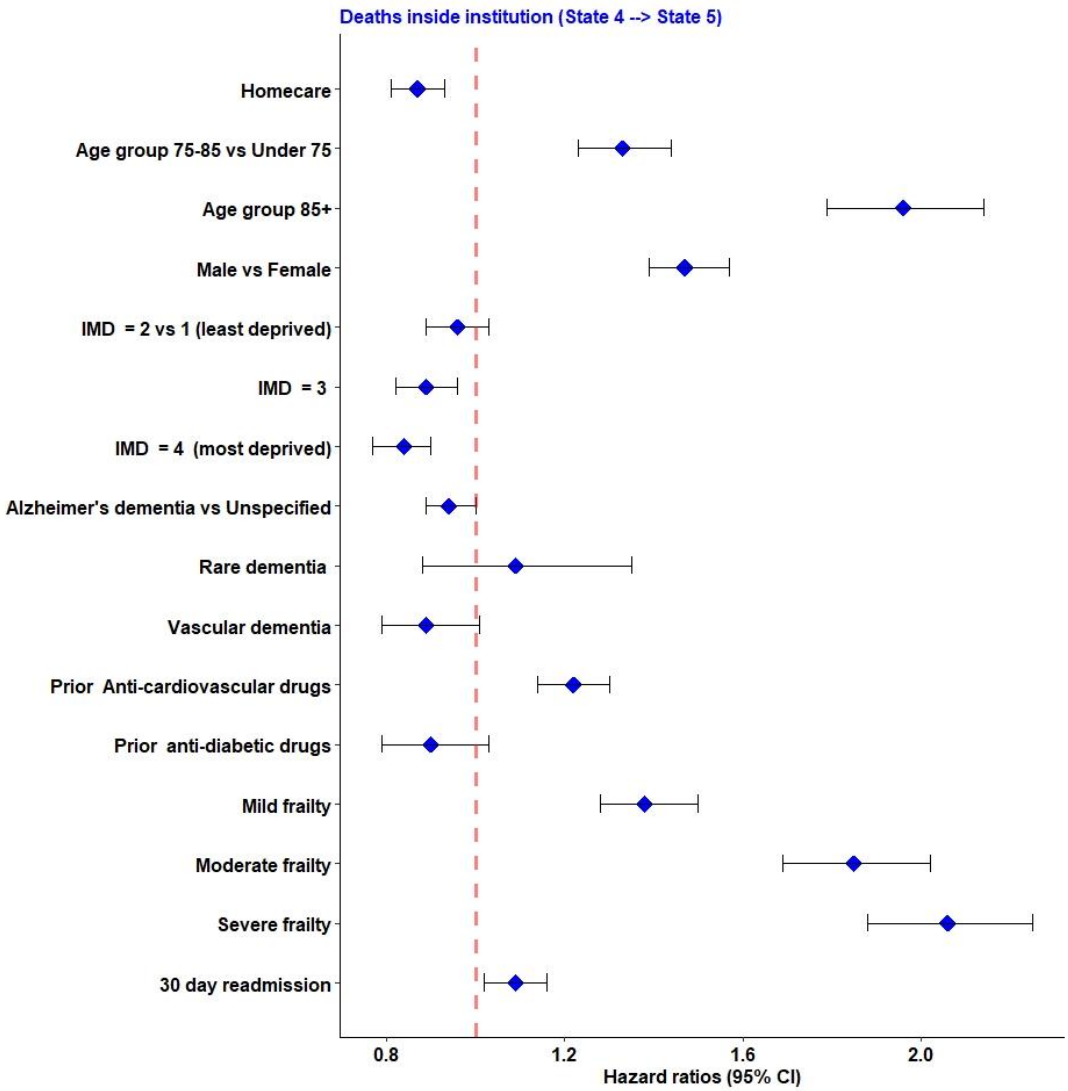


Figure e17: Hazard plot showing the hazard ratios of deaths inside long-term care institutions