Prevalence Measurement for Alzheimer's Disease and Dementia: Current Status and Future Prospects

Julie P.W. Bynum, MD MPH[1]

Kenneth Langa, MD PhD[2]

[1] Department of Medicine, Institute for Healthcare Policy and Innovation, Institute of Gerontology, University of Michigan, Ann Arbor, MI

[2] Department of Medicine, Institute for Social Research, Institute for Healthcare Policy and Innovation, Institute of Gerontology, University of Michigan, and Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI

This paper was commissioned for the Workshop on Challenging Questions about Epidemiology, Care, and Caregiving for People with Alzheimer's Disease and Related Dementias and their Families. The workshop was convened by the Board on Behavioral, Cognitive, and Sensory Sciences on October 17, 2019, in Washington, DC with support from the National Institute on Aging. Opinions and statements included in the paper are solely those of the individual author, and are not necessarily adopted, endorsed, or verified as accurate by the Board on Behavioral, Cognitive, and Sensory Sciences or the National Academy of Sciences, Engineering, and Medicine.

INTRODUCTION

The global population with Alzheimer's Disease and related dementias (AD/ADRD) is expected to nearly triple by 2050. Predictions such as this have important implications for society in general, implying the need for concomitant increases in funding for therapeutic improvements, changes to public policy, and greater focus on the impact of disease on people living with dementia and their caregivers. Accurate, up-to-date information on the societal prevalence of AD/ADRD, and the tracking of potential changes in this prevalence over time have high value for public health and public policy, so close attention to how we obtain these estimates, and potential strategies for improving them in the future, must be a core activity of the scientific community.

In this chapter, we will build toward understanding how to improve measurement of dementia prevalence in the future by making a case for clarity about how we intend to use these data and how well the current data sources fit those intended uses. To do so, we will lay out how the definition of AD/ADRD disease status has evolved, the strengths and limitations of the epidemiological and health care data available for identifying the population that meets these definitions, and the accuracy of current methods of estimating population prevalence. Based on this evaluation, we will describe principles to guide future investments in population assessment strategies for monitoring the population burden of AD/ADRD.

PURPOSE OF POPULATION PREVALENCE MEASURES

"Prevalence" is defined as the proportion of individuals in a defined population with a condition or disease. The use of prevalence measures for any acute or chronic disease falls into

different domains in which health and population needs intersect. In the case of AD/ADRD, the direct use of prevalence data falls into the major domain headings of public health, medical care and public policy. The data collected for these different purposes can be used to quantify disease burden and trends in a population. More specifically, these data are used to study disease etiology, disparities, financial implications, planning strategy and investments, and effectiveness of interventions across each of these domains. But the application dictates the degree of sensitivity, specificity, and other aspects of measurement performance needed, and the added value of obtaining better data given potential additional costs. In other words, different purposes can require different measures across data sources, necessitating nuanced understanding of the constructs being measured.

Population surveillance of disease in the field of epidemiology is a key function of public health. The purpose of public health surveillance is to "assess public health status, track conditions of public health importance, define public health priorities, evaluate public health programs, [and] develop public health research," such as identifying risk factors for disease (Lee et al, 2010). One of the central methods for population surveillance is to use a sampling approach that leads to accurate population estimates by identifying a representative sample; individuals within the sample are then classified as "cases" or "non-cases" using standard and reproducible diagnostic methods. For some diseases, this classification can be done using biological measures, such as blood pressure for the condition of hypertension. In the case of AD/ADRD, which has typically been defined as a syndrome of clinical features (described in more detail below), the approach requires collection of standardized measures of cognition, physical function, and in some approaches a clinical adjudication by experts. Given the purposes for which these data are

to be used, it is essential that classification strategies are *consistent over time* and *accurately reflect disease across geographies and diverse populations*.

Health care, like public health, requires information about the burden of disease and studies of risk factors for disease, but additionally is concerned with how people with the disease access services, which requires knowing the prevalence within their local environment. In addition, medical research often uses epidemiological studies to inform more biologically detailed and etiologic studies regarding the nature of the disease. Observations about prevalence can be used secondarily to focus clinical studies on populations at high risk or alternatively can use comparisons to the general population to identify disparities in treatment or outcomes. For medical care purposes, the key feature of prevalence data is to reflect *the local community* without *bias induced by differences in care seeking*.

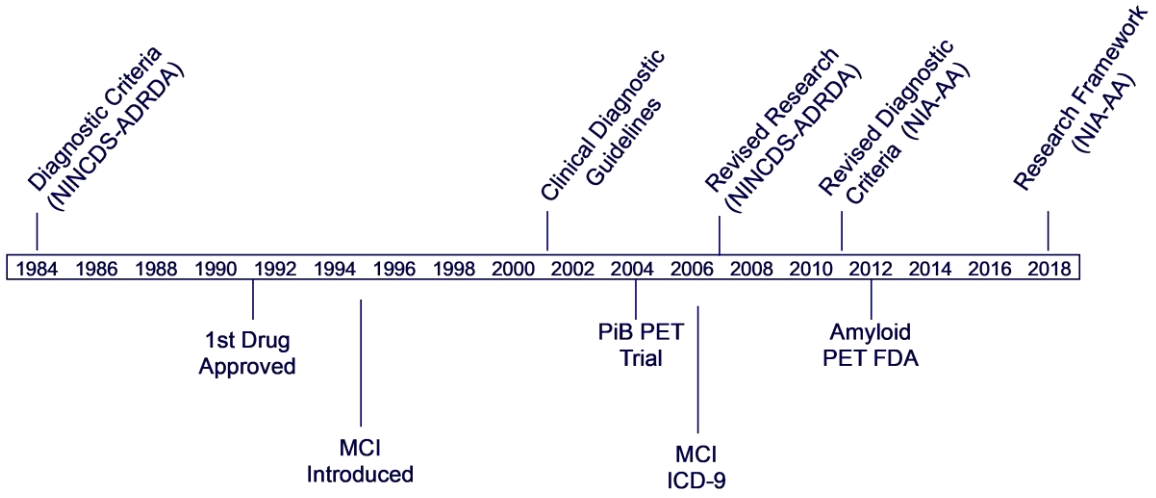
In the public policy domain, interest is focused on anticipating the disease burden and needs of the population looking forward that inform financing of public programs and regulatory actions. Policy makers are concerned with health but are also concerned with issues related to long-term care, labor markets, justice and law enforcement, science and technology, housing and community services and, importantly, how to finance all of those activities. In the case of AD/ADRD there are wide-ranging effects on families and on governments due the prevalence of the disease and the level of services required as the disease progresses; this use of population measures of disease can alter long-term strategy and investments with broad impact. In this context, whatever data approach is taken requires *consensus among the scientific community that it reflects the aspects of disease* that are most important for policy makers to consider as they

forecast and plan for the future. The specific etiology that leads to the syndrome of dementia may be less critical for the majority of governmental activities that address caring for people living with dementia.

DEFINITION OF DISEASE AND EVOLUTION OVER TIME

Over the past few decades, there have been sequential changes in clinical guidelines and diagnostic criteria for defining the presence of AD/ADRD. Figure 1 shows the timing of those important changes, focusing on consensus activities involving the National Institute on Aging.

Figure 1: Timeline for Revisions of Diagnostic Guidelines for AD/ADRD



The diagnostic criteria used to define and identify a disease or condition, as well as the interpretation and operationalization of those criteria in clinical and research settings, necessarily have an important effect on the measured incidence and prevalence of that disease or condition. This is especially true for a geriatric condition such as dementia, given that the onset and progression of cognitive decline is typically gradual and variable in trajectory. In addition, determining when cognitive decline has become severe enough to cause limitations in activities of daily living (ADLs) (a key element of most clinical diagnostic criteria) is challenging, and this

uncertainty regarding when the "dementia threshold" is crossed will likely lead to variation in case definition and prevalence estimates across clinicians and researchers. For instance, an often-cited study of the influence of differing diagnostic criteria on dementia prevalence found a nearly tenfold difference in prevalence in the Canadian Study of Health and Aging when using the ICD-10 criteria (3.1% of those aged 65+) compared to the DSM-III criteria (29.1%) (Erkinjuntti et al, 1997).

In the early 1980s, the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) convened an expert panel to establish criteria for the clinical diagnosis of Alzheimer's disease. The NINCDS-ADRDA criteria (Mckhann et al, 1984) have been widely used in both clinical and research settings. These criteria established the classifications of: Probable Alzheimer's disease; Possible Alzheimer's disease; and Definite Alzheimer's disease, based on evidence from cognitive testing of a decline in two or more cognitive domains that is progressive and has an onset of between age 40 and 90. "Probable AD" requires that there are no other co-occurring conditions that could explain the cognitive decline, while "Possible AD" is diagnosed when there are co-occurring conditions that could explain the cognitive decline, or the clinical course of the decline is atypical. The diagnosis of Probable AD is updated to "Definite AD" when there is biopsy or autopsy evidence of typical AD pathology in the brain. An important distinction to note about the NINCDS-ADRDA criteria is that while limitations in ADLs "support" the diagnosis of "Probable AD," the limitations are not *required* for diagnosis. As noted below, most later diagnostic criteria for dementia due to AD, for example the Diagnostic and Statistical Manual of Mental Disorders Third Edition Revised (DSM-III-R) published in

1987, as well as subsequent DSM updates, require limitations in ADLs for the diagnosis to be made (the dementia "disability criterion"). Whether or not ADL limitations are required for diagnosis will obviously have implications for the measured prevalence of dementia in clinical and research settings.

In 2007, an international working group proposed an update to the 1984 NINCDS-ADRDA criteria that incorporated new imaging and biomarker technologies (e.g., amyloid PET, structural MRI, and cerebrospinal fluid (CSF) markers) as supportive evidence for the diagnosis of AD. Importantly, these criteria identified two early stages of AD, Preclinical AD and Prodromal AD. "Preclinical AD" was defined as an asymptomatic stage in which biomarkers of AD are present, while "Prodromal AD" was defined in a similar way to "Mild Cognitive Impairment," namely the presence of objective declines in cognition that are not yet severe enough to cause significant limitations in ADLs.

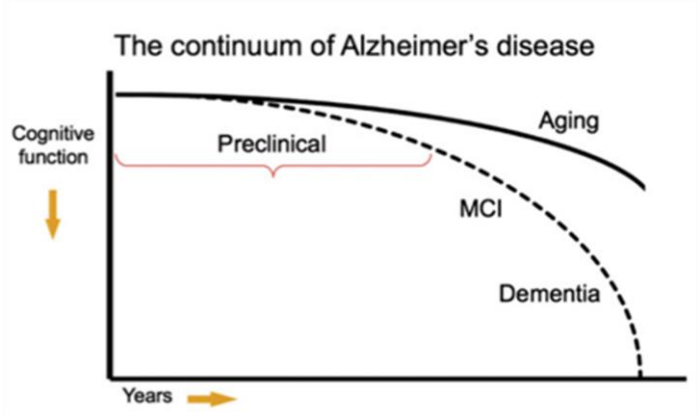
In 2011, the National Institute on Aging and the Alzheimer's Association (NIA-AA) proposed further updates to the 1984 NINCDS-ADRDA diagnostic criteria to incorporate more recent clinical and basic science research regarding the natural history and causes of cognitive decline in adults (Mckhann et al, 2011). The NIA-AA 2011 criteria for dementia are:

1) Impairment in at least two of: memory, reasoning, visuospatial abilities, language; or changes in personality / behavior; and
2) The impairments "interfere with the ability to function at work or at usual activities," and represent a decline from previous levels of functioning; and
3) The impairments are not explained by delirium or psychiatric disorder.

As noted above, the inclusion of the "disability criterion" in the NIA-AA 2011 criteria is likely to lead to a lower prevalence of dementia due to Alzheimer's disease compared to the NINCDS-ADRDA criteria, which did not specifically include the criterion of limitations in work or usual activities.

Like the 2007 international working group criteria, the NIA-AA 2011 criteria designated and defined two additional "pre-dementia" states on the continuum of cognitive and functional decline: 1) Mild Cognitive Impairment (MCI) (Albert et al, 2011); and 2) Preclinical Alzheimer's disease (Sperling et al, 2011) (Figure 2). Like the prodromal AD defined by the 2007 working group, MCI is characterized by subjective and objective declines in cognitive function that are not severe enough to cause disability in daily function (Albert et al, 2011; Langa and Levine, 2014); preclinical AD is characterized by the presence of biomarker and neuroimaging findings of AD pathology in the brain, in the absence of any symptoms of cognitive decline. The identification of MCI and potential shifts over time in the diagnostic "border" between MCI and dementia have important implications for identifying and measuring trends in dementia incidence and prevalence (Matthews et al, 2008), since changes in diagnostic thresholds may lead to differences in measured incidence and prevalence rates over time, separate from any "true" change in the presence of dementia in the population under study.

Figure 2: 2011 Conceptualization of Definitions in NIA-AA 2011 Criteria



The recently-published *NIA-AA Research Framework* that proposes a "biological definition" of AD updates the NIA-AA 2011 concept of preclinical AD, and will likely have an important impact on how the AD continuum is conceptualized, and, by extension, on estimates of incidence and prevalence of AD (Jack et al, 2018). The key change proposed by the 2018 NIA-AA Research Framework is to define AD with three biological measures: Amyloid, Tau, and Neurodegeneration, the AT(N) classification (see Table 1). This AD diagnostic classification using biomarkers is distinct from the clinical symptoms (i.e., measurable cognitive decline that leads to disability in daily life), so under this definition "Alzheimer's disease" now refers to a pathological process, rather than a clinical syndrome (Jack et al, 2018).

The explicit separation of Alzheimer's disease defined biologically from dementia defined clinically has raised a number of concerns regarding the implications for estimating the prevalence of AD in populations (Glymour et al, 2018). One key concern is the difficulty and expense of obtaining the AT(N) biomarkers outside of academic clinical settings. Since the diagnosis of AD using the new NIA-AA research framework is dependent on collection of biomarkers in a clinical setting, it likely will be especially difficult to recruit and retain

individuals from rural communities and from minority populations that are already under-represented in most clinical and epidemiological studies of aging and cognitive decline (Glymour et al, 2018), which will worsen generalizability to rural and minority populations. Conversely, people who have the clinical syndrome of dementia, with progressive impairment of cognition and resultant disability, but none of the currently measured biomarkers for disease may not be counted in estimates of prevalence if one requires biomarker evidence in the disease definition.

Table 1: NIA-AA Research Framework AT(N) Classification System

AT(N) biomarker grouping
A: Aggregated Aβ or associated pathologic state
CSF $A\beta_{42}$, or $A\beta_{42}/A\beta_{40}$ ratio
Amyloid PET
T: Aggregated tau (neurofibrillary tangles) or associated pathologic state
CSF phosphorylated tau
Tau PET
(N): Neurodegeneration or neuronal injury
Anatomic MRI
FDG PET
CSF total tau

Abbreviations: Aβ, β amyloid; CSF, cerebrospinal fluid.

The interest in defining preclinical AD as a stage in which there is biomarker evidence of AD pathology, but no current symptoms of cognitive decline was encouraged by the multiple failures of AD medications for MCI and dementia due to AD, and the belief that treating earlier in the pathological progression (prior to symptom onset) would increase the chances of success. There are now multiple trials of drugs for dementia prevention being tested in individuals with preclinical AD (Karlawish and Langa, 2016).

While there is hope that preclinical AD treatments will lead to prevention or delay of dementia, a number of concerns have been raised regarding the risk of "overdiagnosis" of AD

that may result from the switch to a biomarker-based classification scheme (Langa and Burke, 2019). In this context, overdiagnosis is the identification of AD pathology that doesn't lead to symptomatic dementia, even without treatment. Overdiagnosis exposes individuals to the costs and potential negative side-effects of treatments (as well as potential anxiety and "stigma" associated with the diagnosis), without the benefit of the prevention of clinically significant disease (Langa and Burke, 2019).

One reason that overdiagnosis of preclinical AD could have important consequences is simply the likely large number of individuals who would be diagnosed. Using estimates of AD biomarker prevalence and transition rates from preclinical AD to dementia, Brookmeyer and colleagues (2018) estimated that about 46 million adults had preclinical AD in the US in 2017, compared to a prevalence of dementia due to AD of 3.6 million (Table 2). The shifting conceptualization of AD as a "biological," rather than a "clinical" entity will likely have an important impact on which type of "prevalence" is the focus of future research and policy.

Table 2 : Prevalence (in millions) of preclinical and clinical disease stages of Alzheimer's disease in the United States in 2017 and 2060 based on multistate model [ranges generated by high and low series of transition rates] (Brookmeyer et al, 2018)

	Calendar year	
Disease state	2017	2060
Amyloidosis only [state 2]	22.14 [18.70–26.70]	31.90 [28.04–36.75]
Neurodegeneration only [state 3]	8.33 [5.68–9.16]	13.60 [8.47–16.11]
Amyloidosis and neurodegeneration [state 4]	16.23[11.85–21.93]	30.18 [23.49–37.78]
MCI due to Alzheimer's disease [states 5 + 6]	2.43 [1.41–4.02]	5.70 [3.61–8.34]
With neurodegeneration	0.66 [0.28–1.51]	1.23 [0.56–2.49]
With amyloidosis and neurodegeneration	1.77 [1.13–2.51]	4.47 [3.05–5.85]
Clinical Alzheimer's disease [states 7 + 8]	3.65 [1.70–7.62]	9.30 [4.58–17.82]
Early stage	2.11 [1.03–4.12]	5.29 [2.73–9.40]
Late stage	1.54 [0.67–3.50]	4.01 [1.85–8.42]

CURRENT DATA SOURCES--STRENGTHS AND CHALLENGES

The types of data that can inform prevalence estimates can come from multiple sources, including epidemiological studies and data generated in the course of health care delivery, each with its own strengths and challenges. *Epidemiologic studies* rely on recruiting a sample of the population and determining their cognitive status using a variety of potential methods. The study prevalence is then used to create a population estimate. More recently, data obtained through the provision of medical care have become more routinely available. The cost-effectiveness of using data that are already being collected is attractive. The first of these sources, *administrative claims data*, contains information about diagnosis as documented on the bill submitted for payment by clinicians and health care facilities. The second source is the *electronic health record*. It can contain billing information, but also contains structured medical documentation (such as problem lists and past medical history lists) as well as text of clinical notes. The newest source of information regarding prevalence of disease, or risk for the disease, is *biomarker data* collected as part of epidemiological research or more recently for other clinical studies. Additionally, with the pace of scientific developments, there is the potential to acquire other types of non-invasive or behavioral data that can inform population surveillance.

Epidemiologic data current state

A growing number of population-based surveys of adults in the U.S. and around the world now include measures of cognition that allow determination of level of cognitive function, and, in turn, estimation of dementia incidence and prevalence in the population covered.

The Health and Retirement Study (HRS) and the associated network of international aging surveys patterned on the HRS have been collecting longitudinal cognitive data on middle-aged and older adults since 1992 (Sonnega et al, 2014). The HRS has also embedded more extensive data collections focused on cognitive and physical function in HRS sub-samples with the goal of providing more accurate estimates of dementia incidence and prevalence in its nationally representative sample. The Aging, Demographics, and Memory Study (ADAMS) was fielded in 2001, with follow-up through 2006, and included a 2 to 3-hour in-home assessment of cognition and function in a sub-sample of 856 HRS respondents (Langa et al, 2005). An expert consensus panel used the ADAMS assessment data to assign a dementia diagnosis and its cause. The ADAMS showed that embedding a valid dementia diagnosis in an ongoing nationally representative longitudinal study of older adults creates unique opportunities for identifying the current prevalence and societal impact of MCI (Plassman et al, 2008) and dementia (Plassman et al, 2007), as well as major risk (Iwashyna et al, 2010; Llewellyn et al, 2010) and protective factors (Langa et al, 2008; Langa et al, 2017), and important health (Okura et al, 2010) and economic (Hurd et al, 2013) outcomes. A key limitation of the ADAMS, however, was its high cost per assessment and the resulting small sample size (n=856) that could be assessed due to budgetary limitations. A new HRS dementia sub-study, the Harmonized Cognitive Assessment Protocol (HCAP) project was fielded in 2016 in a more cost-effective manner, using computer-assisted personal interviews administered by specially trained survey interviewers (Langa et al, 2019). The large increase in sample size (n~3,500) will significantly increase the statistical power for analyses using these new data compared to the ADAMS sample, opening up greater opportunities for sub-analyses, such as the relationship of race/ethnicity, education, comorbidities, and geographic location to dementia risk.

In addition to the HRS data collection in the United States, more than 30 other countries around the world have fielded population-based longitudinal studies of aging patterned on the HRS since 2001. The goal of this collaborative network of studies is to obtain comparable data relevant for aging research in order to support international comparative studies that will allow a better understanding of the predictors and outcomes of physical and cognitive decline in aging populations around the globe. Twelve of the HRS international studies have also fielded an HCAP sub-study of dementia in order to facilitate international comparative research on dementia prevalence and outcomes. More information on the HRS and HCAP international studies is available at the HRS website.[1,2]

Administrative data current state

Over the last decades, data created through the process of health care delivery have become more available for use in research. The cost-effectiveness of using data that are already being collected is attractive, especially in the case of diseases that predominantly affect people who are old enough to qualify for Medicare. In addition, the widespread expansion of electronic health records in tandem with developments in the field of data science has made clinical information beyond billing information available.

The strengths of these data come from several features over and above the reduced cost of primary data collection necessary for epidemiologic studies, most clearly demonstrated when

[1] HRS International Sister Studies. Retrieved from https://hrs.isr.umich.edu/about/international-sister-studies

[2] HRS Harmonized Cognitive Assessment Protocol (2016). Retrieved from http://hrsonline.isr.umich.edu/index.php?p=shoavail&iyear=ZU

considering Medicare. The power stems from the nearly complete capture of the population of adults aged 65 and older in the United States through the Medicare system. For this age segment of the U.S. population, the Part A entitlement of Medicare, like Social Security, is for anyone with a 10 year work history or their spouses, which translates to 97% of people age 65 older residing in the U.S.[3] Key subgroups missed could include those who immigrate late in life and partners who do not meet federal spouse eligibility. Another strength of the data is that Medicare requires a diagnosis to pay a claim, so missing diagnosis is infrequent. Finally, Medicare data are centralized with an established process for obtaining access for research purposes.

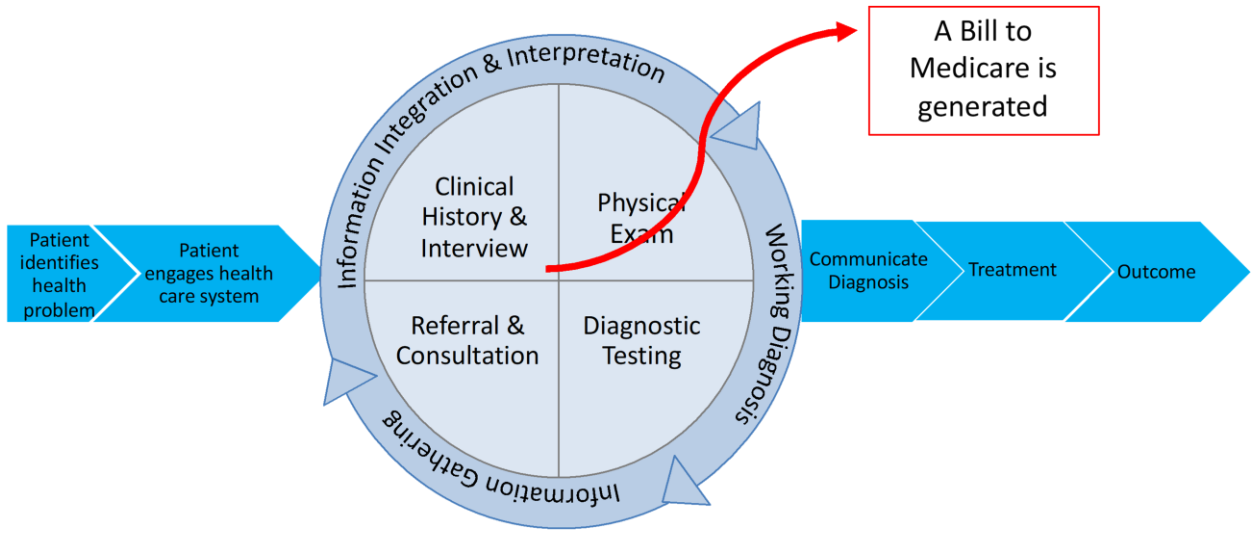
While the population coverage is an obvious strength of these administrative data, there are myriad limitations. The data only include people who are covered by the payer from whom the data is obtained. For people 65 and over and in different payment models (like Medicare Advantage (MA) plans), the rules requiring a diagnosis on a bill for payment differ. For people under 65, data will be housed with insurance payers and with no population-based denominator on which to calculate prevalence. In both MA and commercial insurance for those under age 65, information about why people disenroll (such as changing plan versus dying) is not collected or reported in a standardized way.

The most significant challenge of billing data, however, is the dependence on the clinical

[3] Estimated based in similar eligibility criteria with Social Security Administration and the estimate of SSA coverage for older adult residents of the U.S. (Policy Basics: Top Ten Facts about Social Security. Retrieved from https://www.cbpp.org/research/social-security/policy-basics-top-ten-facts-about-social-security) (Who is Eligible for Medicare? Retrieved from https://www.hhs.gov/answers/medicare-and-medicaid/who-is-elibible-for-medicare/index.html)

process for identifying affected people – a process that can have inherent flaws and biases. The process by which a person receives a diagnosis involves many factors, including those to do with the patient (the person's symptoms, beliefs about those symptoms, and their access to healthcare), the doctors (their general diagnostic skills, their experience with the disease, and the availability of and connections to consultative expertise), the system (the availability of diagnostic technologies and payment for those services), and finally societal influences (the benefits and costs of having the disease label). The National Academy of Medicine (NAM) recently created a conceptualization of the diagnostic process shown in Figure 3 that includes a feedback loop in which later effects (such as treatment effectiveness and outcomes) reinforce the diagnostic process, which is an important point in the context of small or uncertain benefits of current therapies for ADRD.

Figure 2: Process of Diagnosis and Origin of Diagnosis Data in Medicare Claims (Figure adapted from National Academies of Sciences, Engineering, and Medicine, 2015)



This complex process is boiled down to a Medicare bill or a diagnosis placed into a medical record. Clinical data sources, such as administrative billing data, consistently underestimate the prevalence of disease compared to epidemiological studies, reflecting either

under-ascertainment of dementia or under-documentation of the disease. In a study that screened for cognitive impairment but did not intervene, 5 percent of primary care patients over age 60 had moderate to severe cognitive impairment but less than a quarter of those affected had dementia documented as a diagnosis (Boustani et al, 2005). A recent meta-analysis estimates 60 percent of ADRD cases go undetected in the community and in primary care 50 percent are estimated to be undiagnosed across several studies (Lang et al, 2017).

The challenges to making and documenting a diagnosis in AD/ADRD can be summarized into those that occur before, during, or after a health care interaction. In all diseases, financial and physical access to health care providers can be a barrier to care. Financial access is mitigated for many people with concerns about AD/ADRD because the onset of disease is most often after age 65 when Medicare becomes available. Financial access is still likely a barrier for people under age 65 and also people who cannot afford Part B coverage. Another access barrier is availability of providers with expertise in evaluation and treatment of dementia and also transportation to reach those providers. These barriers are structural in nature, but there are also issues of public perception that may reduce care-seeking for cognitive decline. Some people may believe that the changes they are experiencing are normal for their age. In fact, prior studies have found that beliefs about whether cognitive loss is a normal part of aging may contribute to racial disparities in diagnosis (Chin, Negash and Hamilton, 2011; Connell, Scott Roberts and McLaughlin, 2007; Connell et al, 2009). Similarly, disease stigma can dissuade care seeking and has also been suggested to contribute to racial differences (Chin, Negash and Hamilton, 2011; Connell, Scott Roberts and McLaughlin, 2007; Connell et al, 2009).

The process that occurs when a person becomes a patient and interacts with the health care system is less explored in AD/ADRD. Studies of primary care physicians suggest they feel they lack the time and confidence necessary to make an AD/ADRD diagnosis (Boustani et al, 2003; Boustani et al, 2005; Boustani, Schubert & Sennour, 2007). Studies about delayed diagnosis and misdiagnosis exist, but these largely predate the era of electronic health records (Knopman, Donohue, and Gutterman, 2015; Campbell et al, 2014; Mendez, 2006; Reynolds et al, 1988; Roca et al, 1984; Rabins, 1983). They do suggest, however, that misdiagnosis is common, especially in the hospital setting. There is also a large literature about what screening tools are most accurate and about how the Medicare annual wellness visit (which requires cognitive decline screening) has been interpreted (Lin et al, 2013; Cordell et al, 2013; Fowler et al, 2019). The volume of studies, along with U.S. Preventive Services Task Force guidelines that do not recommend screening and now negative trials of screening, likely reflect the lack of consensus about what clinicians should actually do (Moyer, 2014; Fowler et al, 2019). The lack of clarity creates "noise" for practicing doctors who are expected to integrate this information into their workflow. The bottom line is that there is not standardization of how a diagnosis of dementia is made, unlike for example the clear guidelines for assessment of blood pressure for hypertension, and this can lead to misclassification of cases.

Finally, a clinician's documentation of a diagnosis may be influenced by whether she or he intends to action on that diagnosis. For example, when AD medications became more accessible with the introduction of the Part D benefit, treatment rates rose (Fowler et al, 2013); this will tend to increase AD/ADRD prevalence for algorithms that include medication use in identifying cases. The type of physician seen and the physician's beliefs about value of

medications may also affect the likelihood of an AD/ADRD diagnosis detectable in the administrative data, as suggested when a study showed that specialty of the consulting physician (neurology versus geriatrics) influenced whether an anti-dementia drug was started (Koller, Hua, and Bynum, 2016). Another further downstream potential deterrent to care seeking is the possibility that long-term care insurance will not be granted by insurance underwriters if cognitive decline is evident in the medical record (Cornell et al, 2016).

Taken together, the prevalence of disease found in health care data may not reflect true disease prevalence. In many cases, the difference between true disease prevalence and what is observed in another data source can be estimated if the potential biases and confounding are known. But importantly, the types of biases present in any given set of health care data may not be the same across data sources or studies, nor are the potential sources of bias easily observable. For example, public perceptions that influence care seeking may vary by ethnic groups or across regions, and the process of screening including whether the Medicare Annual Wellness Visit is routinely implemented may vary across health systems. In contrast to pre-planned public health surveillance systems, the absence of standardization creates challenges for interpretation that require careful application of health care data when trying to inform population prevalence estimates. In many cases with health care data, the biases can be known and addressed but they have yet to be fully addressed in calculations of population prevalence.

Accuracy of Claims Diagnoses

Given the challenges of administrative claims data, many investigators have examined the accuracy of claims-based diagnoses against other gold- or reference-standards. These studies

have improved our confidence about using claims data to identify people with Alzheimer's dementia. In general, the performance of claims is good at identifying people with dementia, although claims underestimate people with early stage disease as reviewed by St. Germaine (2012). The major weakness is lack of precision of the diagnosis; that is, we cannot be confident of the underlying type of dementia, such as AD versus vascular versus Lewy body. The accuracy of the specific code for AD performs fairly poorly (Taylor et al, 2009). Lewy body, vascular dementia and frontotemporal dementia billing have been assessed, but their validity is largely unknown (Goodman et al, 2013; Drabo et al, 2019). While those specific diagnostic codes exist in both ICD-9 and ICD-10, the vast majority of codes used by clinicians are for non-specific dementia (Goodman et al, 2013; Drabo et al, 2019). This problem is especially acute in ICD-9, which included categories that are not used in modern disease classification systems (e.g. senile dementia, pre-senile dementia).

The second challenge is that often the standard against which the claims data are being tested reflects a sample not representative of the population, such as being drawn from a clinical trial, a subset of payers, a convenience sample, or single area (Newcomer et al, 1999; Albrecht et al, 2018; Taylor et al, 2002; Zhu et al, 2018). For example, Taylor and colleagues (2002) is a validation against Alzheimer's Disease Research Center clinic registries which have a strong volunteer bias and few normal controls (which precludes measurement of specificity). And Zhu and colleagues (2018) is an epidemiological study that samples a population within a relatively small geographic area, which may be strongly influenced by local practice norms and billing habits of the regional clinicians and by the racial and ethnic composition of the population sampled.

Nationally representative epidemiological studies have also been used to validate claims-based diagnoses (ADAMS, HRS). Ideally one would directly compare the performance of different algorithms, however no studies have a national population sample and a gold standard reference, except for ADAMS. ADAMS is sampled to reflect the national population size and demographics of the U.S. population age 71 and older. But that sample frame does not necessarily include variation in billing and clinical practices and it is too small to feel confident that these are captured within the sampled set. ADAMS was used to derive an algorithm based on standardized assessments of cognition and function that are measured in the entire HRS sample (Crimmins et al, 2011). Several different algorithms have been developed that have slightly different performance and target populations (Gianattasio et al, 2019). The original algorithm was then used as a reference standard to determine the accuracy of claims data to classify participants as normal; with cognitive impairment not dementia; and with dementia. It is from this study that the Centers for Medicare and Medicaid Services through its contractor, the Clinical Conditions Warehouse (CCW), created its flag for indicating the presence of AD or ADRD. These flags are currently available to any researcher who purchases the file, which standardizes the way researchers using claims data classify disease.

The CCW algorithm validation was based on ICD-9 coding and was billing practices in 2001. Since that time, we have accessible pharmaceutical data through Part D, we have switched from ICD-9 to ICD-10 in the U.S., and clinical practice and disease definitions have changed, so one might be concerned about the accuracy of the CCW algorithm moving forward. Recently, investigators are generating studies evaluating which diagnoses are being used in Medicare claims in the more recent paper (Goodman et al, 2017; Thunell, Freido, and Zissimopoulos,

2019), but a validation of the CCW algorithm using contemporary billing data is not yet published. Investigators are or will be working across multiple institutions to validate their algorithms in other data sets, many of which have now or will soon be linked to Medicare claims. The increased analysis is welcome, yet will require subsequent activities to align and harmonize these algorithms.

Increasingly, nationally or otherwise sampled studies have linked their study cohorts to Medicare administrative data. It is from these linkages that we have seen studies validating the performance of claims. These types of studies are likely to continue to appear as each study independently tests their identified cases against claims. The challenge that arises from these studies is that no single study will be able to state the precision of the claims, because each will represent the clinical practices of the clinical provider practice represented by the sample. Yet if taken together, we may be able to generate a more comprehensive understanding of the relationship of disease presence and appearance in claims by studying performance across studies which sample different populations. Currently multiple epidemiological studies have either already linked to claims (such as National Health and Aging Trends Study (NHATS), Washington Heights-Inwood Columbia Aging Project (WHICAP)) or are in the process of doing so (Appendix Table 2).

Studies that Combine Data Sources

One of the newer developments in population studies is combining multiple of the data types already discussed within a single study. Objective measures of cognition have been combined with biomarker studies, so allowing more detailed knowledge of how the syndrome of dementia relates to the measures observed in imaging, CSF and pathology. Due to the complexity of recruitment and cost of these studies, they often are limited to regional or

convenience samples rather than national samples, such as WHICAP and the Religious Orders Study. Studies such as HRS use population sampling methods and collect survey-based measures on large samples, which are then supplemented with sub-studies that collect clinical cognitive status exams (ADAMS) and more recently neuroimaging.

Given the discussion of changing diagnostic criteria for AD that include both clinical and biological definitions of disease, it is important to note that there is currently no nationally representative sample that has all of the data types one would need to capture the current research and clinical diagnostic criteria. Such a data source would require a national sample, survey-based and/or in-person assessments of cognitive and functional status, as well as neuroimaging or CSF studies. While HRS is beginning to collect imaging data with a small sub-set of its participants, doing so on a scale to allow robust national estimates of change over time would be extremely costly.

NEW DEVELOPMENTS IN POPULATION PREVALENCE MEASUREMENT

Rise of new technologies for cognition and biomarker assessment

The pace of innovation and development in both tools for assessing cognition or risk of cognitive decline and identification of new biomarkers makes for potential major advances in resolving the challenge of obtaining large population data with little burden to participants and lower costs for investigators. For example, use of handwriting analysis collected through a tablet could increase the precision of categorizing people with very early symptoms. Currently accepted biomarkers require amyloid imaging of the brain that necessitates infusion and MRI scanning or lumbar puncture to obtain CSF. Both of these pose some risks to participants and limit recruitment, particularly in some ethnic and racial populations. If newer studies use

peripheral blood or retinal scans for amyloid, for example, population studies will be able to close the gap in our ability to assess underlying risk in the population based on biomarkers. One of the challenges of these new technologies is that it is difficult to anticipate which of these will hold up against the test of further scientific discovery and be useful long into the future once inserted into large, longitudinal population-based studies.

Electronic Health Record Data

Recently, publications are emerging that seek to identify people with AD/ADRD using electronic health record (EHR) data. These data are complex in that they can include multiple different types of data contained within a health systems data warehouse, including the billing data, structure data elements that list diagnosis, and increasingly the text within notes. Like claims data, EHR data have the same limitation of relying on presenting to the health care system to be identified, although EHR data may be more sensitive with the added data elements. Among all these studies few have validated their identified population against a gold or even a reference standard that includes objective cognitive measurement. Most use the structured diagnosis data from the billing process in EHRs or use case review to confirm diagnosis rather than compare to determination of disease status using standardized clinical assessments (Shao et al, 2019; Reuben et al, 2017; McGuiness et al, 2019).

NON-AD FORMS OF DEMENTIA

An important contribution of population-based studies of dementia over the past 25 years is the clear evidence that multiple neuropathological changes are typically present in the brains of older adults, so it is rare for there to be a single pathological cause for dementia. In other words, nearly all dementia in representative samples of older adults is “mixed dementia,” rather

than “pure” AD, vascular dementia, or frontotemporal dementia. For instance, Boyle and colleagues (2018) showed that, in the Religious Orders Study (ROS)/Memory and Aging Project (MAP) population-based samples, nearly 60% of those who died had three or more neuropathologies (including AD; Lewy body; hippocampal sclerosis; TDP 43; vascular infarcts; arteriolar sclerosis; atherosclerosis; cerebral amyloid angiopathy), and there were more than 230 different neuropathologic combinations, each of which occurred in less than 6% of the cohort. In addition, in the ROS/MAP sample, currently identified neuropathologies accounted for only about 40 percent of the variation in late-life cognitive decline, suggesting that there are still other pathologies or causal pathways yet to be identified that have a significant impact on the risk for dementia in later life (Boyle et al, 2013).

The complexity of the neuropathological causes for typical late-life dementia among older adults raises the question of whether it is feasible for population-based studies to make valid estimates of the different sub-types of dementia, and whether the significant cost associated with neuroimaging (as in the case of vascular dementia), neurological exams (to distinguish Lewy Body dementia from Parkinson’s with dementia) and collection of neuropathological data in large population-based studies would be a good scientific investment. In addition to the need for additional testing, the sampling frames and size may have to be large to reach important sub-populations. For example, the younger age of onset of frontotemporal dementia may necessitate larger samples of people aged 45-60.

SUMMARY OF THE CHALLENGES AND RECOMMENDATIONS

Accurate measurement of population prevalence of AD/ADRD is a critical activity for social and behavioral sciences related to aging. Its importance has been recognized by the

initiation of the National Alzheimer's Plan[4] and significant increase in public funding to address the burden these diseases place on individuals, families, and society. Over decades of measuring prevalence, we have gained much knowledge and also identified challenges in our ability to assess population prevalence with accuracy in a way that can keep pace with evolving scientific conceptualization of the disease status.

One can consider each of the data sources discussed above in terms of its validity for estimating the prevalence of true disease, but given the shifting scientific definitions of what it means to have dementia due to Alzheimer's disease, it is also critical to consider a data source's suitability for the specific purpose of the measured prevalence, or *fitness for use* (Weiskopf et al, 2017). The term "fitness for use" comes out of the electronic health record secondary use literature but has applicability to use of data in other contexts. One of the key principles to guide investments in future measurement is to acknowledge that there can be no single best source of data, and that the value of each data source depends on how the information gleaned from the data will be put to use. We propose that to answer the question of how future data collection to assess population prevalence of AD/ADRD should proceed requires a nuanced understanding of the assumptions, biases, and inherent limitations of any given method when no single data source can be perfect.

Since the impact of cognitive decline on individuals, families, and public programs becomes most prominent when the disability "threshold" is reached, we believe a key priority for future dementia prevalence measurement is to continue to track the "dementia syndrome"

[4] National Alzheimer's Plan. Retrieved from https://aspe.hhs.gov/national-plan-address-alzheimers-disease-and-other-napa-documents

characterized by both cognitive and functional impairment. This definition has the benefit of consistency over time, even though it remains a challenge to implement consistently across populations with differing education, race / ethnicity or other social factors. Efforts to harmonize measures of dementia prevalence across studies are as critically important as they were when noted by Wilson and colleagues (2011) nearly a decade ago.

While large-scale population-based studies to provide estimates of dementia prevalence will provide vital information on the current and future trends in the societal burden of the disability related to cognitive decline, such studies would, due to challenges related to subject recruitment and high cost, typically lack the "deep phenotyping" of neuroimaging and CSF biomarkers that are required for the biological AT(N) definition of Alzheimer's disease proposed in the recent NIA-AA Research Framework (Jack et al, 2018). Hopefully, there will be significant advances in biomarker assessments (e.g., blood-based biomarkers) that are less invasive and less costly that would make collection in large representative samples much more feasible.

The importance and value of accurate estimates of the prevalence of biologically defined AD in individuals without symptoms (i.e., preclinical AD) would increase significantly if and when successful interventions targeted at the preclinical stage of AD are identified; at that time, investing additional resources in large-scale population-based biomarker assessments to establish the prevalence of preclinical AD would provide important information on the size of the population that would benefit from those treatments. Since the prevalence of biologically defined Alzheimer's disease will be much larger, perhaps ten times larger (Brookmeyer et al., 2018), than the prevalence of dementia defined using cognitive and functional criteria, it will be

extremely important to clearly define the difference between the two disease definitions for patients, the general public, and policy makers (Karlawish et al., 2017).

Large-scale studies to identify dementia prevalence based on cognitive and functional impairment will also typically lack the neuroimaging and other clinical data required to accurately assess the prevalence of different dementia sub-types or etiologies (e.g., AD, vascular dementia, or frontotemporal dementia). To better understand the population burden of these different dementia sub-types, as well as better understand potential differences in risk factors or outcomes of the sub-types, smaller scale studies that collect the required clinical, imaging, and biomarker data to accurately estimate the prevalence of sub-types would be necessary (Launer 2011).

Diversifying representation across all socioeconomic, racial and ethnic, gender, geographic and health groups is of high priority. This is particularly true as we see differing estimates across studies for African-Americans (Barnes and Bennett, 2014) and as changing educational level is implicated in declining prevalence of AD (Satizabal et al, 2016). The need for better representation of populations currently under-represented in dementia studies will be especially challenging if studies require collection of biomarker data that can only be collected in academic clinical settings. As noted above, development of strategies for more cost-effective collection of imaging and biomarker data in representative population-based samples will be important if accurate estimates of current and future prevalence of preclinical AD are to be available.

Summary

Fundamentally, across the critical uses of population prevalence measurement of AD/ADRD in public health, health care, and public policy, the data approach needs to rest on clear and agreed upon measures and representation that are consistent over time. That being said, there are specialized uses within each domain that may not be served by a unitary approach to data collection. The challenge is to use approaches to meet the needs of end-users across domains that can still be cross-walked and understood in the context of other methods. To the degree that different approaches give differing estimates, we must be able to understand why they are different, employ the measures appropriately for their use case, and articulate clearly what our measures mean when speaking with the broad community of stakeholders.

Reference List

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Albrecht, J. S., Hanna, M., Kim, D., & Perfetto, E. M. (2018). Predicting Diagnosis of Alzheimer's Disease and Related Dementias Using Administrative Claims. *J Manag Care Spec Pharm, 24*(11), 1138-1145. doi:10.18553/jmcp.2018.24.11.1138

Barnes, L. L., & Bennett, D. A. (2014). Alzheimer's disease in African Americans: risk factors and challenges for the future. *Health affairs (Project Hope), 33*(4), 580-586.

Bennett, D. A., Schneider, J. A., Arvanitakis, Z., & Wilson, R. S. (2012). Overview and findings from the religious orders study. *Curr Alzheimer Res, 9*(6), 628-645. doi:10.2174/156720512801322573

Borson, S., Scanlan, J. M., Watanabe, J., Tu, S. P., & Lessig, M. (2006). Improving identification of cognitive impairment in primary care. *Int J Geriatr Psychiatry, 21*(4), 349-355. doi:10.1002/gps.1470

Boustani, M., Callahan, C. M., Unverzagt, F. W., Austrom, M. G., Perkins, A. J., Fultz, B. A., . . . Hendrie, H. C. (2005). Implementing a Screening and Diagnosis Program for Dementia in Primary Care. *Journal of General Internal Medicine, 20*(7), 572-577. doi:10.1111/j.1525-1497.2005.0126.x

Boustani, M., Peterson, B., Hanson, L., Harris, R., & Lohr, K. N. (2003). Screening for dementia in primary care: A summary of the evidence for the u.s. preventive services task force. *Annals of Internal Medicine, 138*(11), 927-937. doi:10.7326/0003-4819-138-11-200306030-00015

Boustani, M., Schubert, C., & Sennour, Y. (2007). The challenge of supporting care for dementia in primary care. *Clinical Interventions in Aging, 2*(4), 631-636.

Boyle, P. A., Wilson, R. S., Yu, L., Barr, A. M., Honer, W. G., Schneider, J. A., & Bennett, D. A. (2013). Much of late life cognitive decline is not due to common neurodegenerative pathologies. *Ann Neurol, 74*(3), 478-489. doi:10.1002/ana.23964

Boyle, P. A., Yu, L., Wilson, R. S., Leurgans, S. E., Schneider, J. A., & Bennett, D. A. (2018). Person-specific contribution of neuropathologies to cognitive loss in old age. *Ann Neurol, 83*(1), 74-83. doi:10.1002/ana.25123

Brookmeyer, R., Abdalla, N., Kawas, C. H., & Corrada, M. M. (2018). Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. *Alzheimers Dement, 14*(2), 121-129. doi:10.1016/j.jalz.2017.10.009

Callahan, C. M., Hendrie, H. C., & Tierney, W. M. (1995). Documentation and evaluation of cognitive impairment in elderly primary care patients. *Ann Intern Med, 122*(6), 422-429. doi:10.7326/0003-4819-122-6-199503150-00004

Campbell, N. L., Cantor, B. B., Hui, S. L., Perkins, A., Khan, B. A., Farber, M. O., . . . Boustani, M. A. (2014). Race and documentation of cognitive impairment in hospitalized older adults. *J Am Geriatr Soc, 62*(3), 506-511. doi:10.1111/jgs.12691

Chin, A. L., Negash, S., & Hamilton, R. (2011). Diversity and disparity in dementia: the impact of ethnoracial differences in Alzheimer disease. *Alzheimer disease and associated disorders, 25*(3), 187-195.

Connell, C. M., Scott Roberts, J., & McLaughlin, S. J. (2007). Public opinion about Alzheimer disease among blacks, hispanics, and whites: results from a national survey. *Alzheimer disease and associated disorders, 21*(3), 232-240.

Connell, C. M., Scott Roberts, J., McLaughlin, S. J., & Akinleye, D. (2009). Racial differences in knowledge and beliefs about Alzheimer disease. *Alzheimer disease and associated disorders, 23*(2), 110-116.

Cordell, C. B., Borson, S., Boustani, M., Chodosh, J., Reuben, D., Verghese, J., . . . Medicare Detection of Cognitive Impairment, W. (2013). Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimers Dement, 9*(2), 141-150. doi:10.1016/j.jalz.2012.09.011

Cornell, P. Y., Grabowski, D. C., Cohen, M., Shi, X., & Stevenson, D. G. (2016). Medical Underwriting In Long-Term Care Insurance: Market Conditions Limit Options For Higher-Risk Consumers. *Health Aff (Millwood), 35*(8), 1494-1503. doi:10.1377/hlthaff.2015.1133

Crimmins, E. M., Kim, J. K., Langa, K. M., & Weir, D. R. (2011). Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. *J Gerontol B Psychol Sci Soc Sci, 66 Suppl 1*, i162-171. doi:10.1093/geronb/gbr048

Drabo, E. F., Barthold, D., Joyce, G., Ferido, P., Chang Chui, H., & Zissimopoulos, J. (2019). Longitudinal analysis of dementia diagnosis and specialty care among racially diverse Medicare beneficiaries. *Alzheimers Dement, 15*(11), 1402-1411. doi:10.1016/j.jalz.2019.07.005

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., . . . Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol, 6*(8), 734-746. doi:10.1016/S1474-4422(07)70178-3

Erkinjuntti, T., Ostbye, T., Steenhuis, R., & Hachinski, V. (1997). The effect of different diagnostic criteria on the prevalence of dementia. *N Engl J Med, 337*(23), 1667-1674. doi:10.1056/NEJM199712043372306

Fowler, N. R., Campbell, N. L., Pohl, G. M., Munsie, L. M., Kirson, N. Y., Desai, U., . . . Boustani, M. A. (2018). One-Year Effect of the Medicare Annual Wellness Visit on Detection of Cognitive Impairment: A Cohort Study. *J Am Geriatr Soc, 66*(5), 969-975. doi:10.1111/jgs.15330

Fowler, N. R., Chen, Y. F., Thurton, C. A., Men, A., Rodriguez, E. G., & Donohue, J. M. (2013). The impact of Medicare prescription drug coverage on the use of antidementia drugs. *BMC Geriatr, 13*, 37. doi:10.1186/1471-2318-13-37

Fowler, N. R., Perkins, A. J., Gao, S., Sachs, G. A., & Boustani, M. A. (2019). Risks and Benefits of Screening for Dementia in Primary Care: The Indiana University Cognitive Health Outcomes Investigation of the Comparative Effectiveness of Dementia Screening (IU CHOICE)Trial. *J Am Geriatr Soc*. doi:10.1111/jgs.16247

Gianattasio, K. Z., Wu, Q., Glymour, M. M., & Power, M. C. (2019). Comparison of Methods for Algorithmic Classification of Dementia Status in the Health and Retirement Study. *Epidemiology, 30*(2), 291-302. doi:10.1097/ede.0000000000000945

Glymour, M. M., Brickman, A. M., Kivimaki, M., Mayeda, E. R., Chene, G., Dufouil, C., & Manly, J. J. (2018). Will biomarker-based diagnosis of Alzheimer's disease maximize

scientific progress? Evaluating proposed diagnostic criteria. *Eur J Epidemiol, 33*(7), 607-612. doi:10.1007/s10654-018-0418-4

Goodman, R. A., Lochner, K. A., Thambisetty, M., Wingo, T. S., Posner, S. F., & Ling, S. M. (2017). Prevalence of dementia subtypes in United States Medicare fee-for-service beneficiaries, 2011-2013. *Alzheimers Dement, 13*(1), 28-37. doi:10.1016/j.jalz.2016.04.002

Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J., & Langa, K. M. (2013). Monetary costs of dementia in the United States. *N Engl J Med, 368*(14), 1326-1334. doi:10.1056/NEJMsa1204629

Iliffe, S., Robinson, L., Brayne, C., Goodman, C., Rait, G., Manthorpe, J., & Ashley, P. (2009). Primary care and dementia: 1. diagnosis, screening and disclosure. *International journal of geriatric psychiatry, 24*(9), 895-901.

Iwashyna, T. J., Ely, E. W., Smith, D. M., & Langa, K. M. (2010). Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA, 304*(16), 1787-1794. doi:10.1001/jama.2010.1553

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Karlawish, J., Jack, C. R., Jr., Rocca, W. A., Snyder, H. M., & Carrillo, M. C. (2017). Alzheimer's disease: The next frontier-Special Report 2017. *Alzheimers Dement, 13*(4), 374-380. doi:10.1016/j.jalz.2017.02.006

Karlawish, J., & Langa, K. M. (2016). Unfinished Business in Preventing Alzheimer Disease. *JAMA Intern Med, 176*(12), 1739-1740. doi:10.1001/jamainternmed.2016.6310

Knopman, D., Donohue, J. A., & Gutterman, E. M. (2000). Patterns of care in the early stages of Alzheimer's disease: impediments to timely diagnosis. *J Am Geriatr Soc, 48*(3), 300-304. doi:10.1111/j.1532-5415.2000.tb02650.x

Koller, D., Hua, T., & Bynum, J. P. (2016). Treatment Patterns with Antidementia Drugs in the United States: Medicare Cohort Study. *J Am Geriatr Soc, 64*(8), 1540-1548. doi:10.1111/jgs.14226

Lang, L., Clifford, A., Wei, L., Zhang, D., Leung, D., Augustine, G., . . . Chen, R. (2017). Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open, 7*(2), e011146. doi:10.1136/bmjopen-2016-011146

Langa, K. M., & Burke, J. F. (2019). Preclinical Alzheimer Disease-Early Diagnosis or Overdiagnosis? *JAMA Intern Med*. doi:10.1001/jamainternmed.2019.2629

Langa, K. M., Larson, E. B., Crimmins, E. M., Faul, J. D., Levine, D. A., Kabeto, M. U., & Weir, D. R. (2017). A Comparison of the Prevalence of Dementia in the United States in 2000 and 2012. *JAMA Intern Med, 177*(1), 51-58. doi:10.1001/jamainternmed.2016.6807

Langa, K. M., Larson, E. B., Karlawish, J. H., Cutler, D. M., Kabeto, M. U., Kim, S. Y., & Rosen, A. B. (2008). Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity? *Alzheimers Dement, 4*(2), 134-144. doi:10.1016/j.jalz.2008.01.001

Langa, K. M., & Levine, D. A. (2014). The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA, 312*(23), 2551-2561. doi:10.1001/jama.2014.13806

Langa, K. M., Plassman, B. L., Wallace, R. B., Herzog, A. R., Heeringa, S. G., Ofstedal, M. B., . . . Willis, R. J. (2005). The Aging, Demographics, and Memory Study: study design and methods. *Neuroepidemiology, 25*(4), 181-191. doi:10.1159/000087448

Langa, K. M., Ryan, L. H., McCammon, R. J., Jones, R. N., Manly, J. J., Levine, D. A., . . . Weir, D. R. (2019). The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study Design and Methods. *Neuroepidemiology*, 1-11. doi:10.1159/000503004

Launer, L. J. (2011). Counting dementia: There is no one "best" way. *Alzheimers Dement, 7*(1), 10-14. doi:10.1016/j.jalz.2010.11.003

Lee, L. M., Teutsch, S. M., Thacker, S. B., & St.Louis, M. E. (2010). Principles and Practice of Public Health Surveillance In L. Lee, S. Teutsch, S. Thacker, & M. St.Louis (Eds.), *Principles and Practice of Public Health Surveillance*. New York: Oxford University Press.

Lin, J. S., O'Connor, E., Rossom, R. C., Perdue, L. A., & Eckstrom, E. (2013). Screening for cognitive impairment in older adults: A systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med, 159*(9), 601-612. doi:10.7326/0003-4819-159-9-201311050-00730

Llewellyn, D. J., Lang, I. A., Matthews, F. E., Plassman, B. L., Rogers, M. A., Morgenstern, L. B., . . . Langa, K. M. (2010). Vascular health, diabetes, APOE and dementia: the Aging, Demographics, and Memory Study. *Alzheimers Res Ther, 2*(3), 19. doi:10.1186/alzrt43

McGuinness, L. A., Warren-Gash, C., Moorhouse, L. R., & Thomas, S. L. (2019). The validity of dementia diagnoses in routinely collected electronic health records in the United Kingdom: A systematic review. *Pharmacoepidemiol Drug Saf, 28*(2), 244-255. doi:10.1002/pds.4669

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*(7), 939-944. doi:10.1212/wnl.34.7.939

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., . . . Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 263-269. doi:10.1016/j.jalz.2011.03.005

Mendez, M. F. (2006). The accurate diagnosis of early-onset dementia. *Int J Psychiatry Med, 36*(4), 401-412. doi:10.2190/Q6J4-R143-P630-KW41

Moyer, V. A. (2014). Screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med, 160*(11), 791-797. doi:10.7326/m14-0496

National Academies of Sciences, Engineering, and Medicine. 2015. *Improving diagnosis in health care*. Washington, DC: The National Academies Press.

Newcomer, R., Clay, T., Luxenberg, J. S., & Miller, R. H. (1999). Misclassification and selection bias when identifying Alzheimer's disease solely from Medicare claims records. *J Am Geriatr Soc, 47*(2), 215-219.

Okura, T., Plassman, B. L., Steffens, D. C., Llewellyn, D. J., Potter, G. G., & Langa, K. M. (2010). Prevalence of neuropsychiatric symptoms and their association with functional

limitations in older adults in the United States: the aging, demographics, and memory study. *J Am Geriatr Soc, 58*(2), 330-337. doi:10.1111/j.1532-5415.2009.02680.x

Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., . . . Rodgers, W. L. (2007). Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology, 29*(1-2), 125-132.

Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., . . . Wallace, R. B. (2008). Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med, 148*(6), 427-434.

Ponjoan, A., Garre-Olmo, J., Blanch, J., Fages, E., Alves-Cabratosa, L., Marti-Lluch, R., . . . Ramos, R. (2019). Epidemiology of dementia: prevalence and incidence estimates using validated electronic health records from primary care. *Clin Epidemiol, 11*, 217-228. doi:10.2147/CLEP.S186590

Rabins, P. V. (1983). Reversible dementia and the misdiagnosis of dementia: a review. *Hosp Community Psychiatry, 34*(9), 830-835. doi:10.1176/ps.34.9.830

Reuben, D., Hackbarth, A., Wenger, N., Tan, Z., & Jennings, L. (2017). An Automated Approach to Identifying Patients with Dementia Using Electronic Medical Records. *JAGS, 65*(3), 658-659.

Reynolds, C. F., 3rd, Hoch, C. C., Kupfer, D. J., Buysse, D. J., Houck, P. R., Stack, J. A., & Campbell, D. W. (1988). Bedside differentiation of depressive pseudodementia from dementia. *Am J Psychiatry, 145*(9), 1099-1103. doi:10.1176/ajp.145.9.1099

Roca, R. P., Klein, L. E., Kirby, S. M., McArthur, J. C., Vogelsang, G. B., Folstein, M. F., & Smith, C. R. (1984). Recognition of dementia among medical patients. *Arch Intern Med, 144*(1), 73-75.

Satizabal, C. L., Beiser, A. S., Chouraki, V., Chene, G., Dufouil, C., & Seshadri, S. (2016). Incidence of Dementia over Three Decades in the Framingham Heart Study. *N Engl J Med, 374*(6), 523-532. doi:10.1056/NEJMoa1504327

Shao, Y., Zeng, Q. T., Chen, K. K., Shutes-David, A., Thielke, S. M., & Tsuang, D. W. (2019). Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records. *BMC Med Inform Decis Mak, 19*(1), 128. doi:10.1186/s12911-019-0846-4

Sonnega, A., Faul, J. D., Ofstedal, M. B., Langa, K. M., Phillips, J. W., & Weir, D. R. (2014). Cohort Profile: the Health and Retirement Study (HRS). *Int J Epidemiol, 43*(2), 576-585. doi:10.1093/ije/dyu067

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., . . . Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 280-292. doi:10.1016/j.jalz.2011.03.003

St Germaine-Smith, C., Metcalfe, A., Pringsheim, T., Roberts, J. I., Beck, C. A., Hemmelgarn, B. R., . . . Jette, N. (2012). Recommendations for optimal ICD codes to study neurologic conditions: a systematic review. *Neurology, 79*(10), 1049-1055. doi:10.1212/WNL.0b013e3182684707

Taylor, D. H., Jr., Fillenbaum, G. G., & Ezell, M. E. (2002). The accuracy of medicare claims data in identifying Alzheimer's disease. *J Clin Epidemiol, 55*(9), 929-937.

Taylor, D. H., Jr., Ostbye, T., Langa, K. M., Weir, D., & Plassman, B. L. (2009). The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. *J Alzheimers Dis, 17*(4), 807-815. doi:10.3233/JAD-2009-1099

Thunell, J., Ferido, P., & Zissimopoulos, J. (2019). Measuring Alzheimer's Disease and Other Dementias in Diverse Populations Using Medicare Claims Data. *J Alzheimers Dis, 72*(1), 29-33. doi:10.3233/JAD-190310

Valcour, V. G., Masaki, K. H., Curb, J. D., & Blanchette, P. L. (2000). The detection of dementia in the primary care setting. *Arch Intern Med, 160*(19), 2964-2968. doi:10.1001/archinte.160.19.2964

Weiskopf, N., Bakken, S., George, H., & Weng, C. (2017). A Data Quality Assessment Guideline for Electronic Health Record Data Reuse. *EGEMS (Wash DC), 5*(1). doi:10.13063/egems.1280.s1

Wilkins, C. H., Wilkins, K. L., Meisel, M., Depke, M., Williams, J., & Edwards, D. F. (2007). Dementia undiagnosed in poor older adults with functional impairment. *J Am Geriatr Soc, 55*(11), 1771-1776. doi:10.1111/j.1532-5415.2007.01417.x

Wilkinson, T., Ly, A., Schnier, C., Rannikmae, K., Bush, K., Brayne, C., . . . Dementias Platform, U. K. (2018). Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimers Dement, 14*(8), 1038-1051. doi:10.1016/j.jalz.2018.02.016

Wilson, R. S., Weir, D. R., Leurgans, S. E., Evans, D. A., Hebert, L. E., Langa, K. M., . . . Bennett, D. A. (2011). Sources of variability in estimates of the prevalence of Alzheimer's disease in the United States. *Alzheimers Dement, 7*(1), 74-79. doi:10.1016/j.jalz.2010.11.006

Zhu, C. W., Ornstein, K. A., Cosentino, S., Gu, Y., Andrews, H., & Stern, Y. (2019). Misidentification of Dementia in Medicare Claims and Related Costs. *J Am Geriatr Soc, 67*(2), 269-276. doi:10.1111/jgs.15638