ELSEVIER

Alzheimer's & Dementia 6 (2010) 334–341

Alzheimer's & Dementia

# An examination of Alzheimer's disease case definitions using Medicare claims and survey data

Pei-Jung Lin[a,*], Daniel I. Kaufer[b], Matthew L. Maciejewski[c,d], Rahul Ganguly[e], John E. Paul[f], Andrea K. Biddle[f]

[a]Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA
[b]Department of Neurology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA
[c]Center for Health Services Research in Primary Care, Durham VA Medical Center, Durham, NC, USA
[d]Division of General Internal Medicine, Department of Medicine, Duke University Medical Center, Durham, NC, USA
[e]Global Health Outcomes, GlaxoSmithKline, Research Triangle Park, NC, USA
[f]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

**Abstract**

**Background:** The prevalence and expenditure estimates of Alzheimer's disease (AD) from studies using one data source to define cases vary widely. The objectives of this study were to assess agreement between AD case definitions classified with Medicare claims and survey data and to provide insight into causes of widely varied expenditure estimates.

**Methods:** Data were obtained from the 1999–2004 Medicare Current Beneficiary Survey linked with Medicare claims (n = 57,669). Individuals with AD were identified by survey, diagnosis, use of an AD prescription medicine, or some combination thereof. We also explored how much health care and drug expenditures vary by AD case definition.

**Results:** The prevalence of AD differed significantly by case definition. Using survey report alone yielded more cases (n = 1,994 or 3.46%) than diagnosis codes alone (n = 1,589 or 2.76%) or Alzheimer's medication use alone (n = 1,160 or 2.01%). Agreement between case definitions was low, with kappa coefficients ranging from 0.37 to 0.40. Per capita health expenditures ranged from $16,547 to $24,937, and drug expenditures ranged from $2,303 to $3,519, depending on how AD was defined.

**Conclusions:** Different information sources yield widely varied prevalence and expenditure estimates. Although claims data provided a more objective means for identifying AD cases, survey report identified more cases, and pharmacy data also are an important source for case ascertainment. Using any single source will underestimate the prevalence and associated cost of AD. The wide range of AD cases identified by using different data sources demands caution interpreting cost-of-illness studies using single data sources.



*Keywords:* Alzheimer's disease; Epidemiologic methods; Medicare; Prevalence; Expenditure

## 1. Introduction

Alzheimer's disease (AD) is the most common form of dementia in the elderly, comprising approximately half of all dementia cases [1]. Great variation exists in estimates of current and projected AD prevalence derived from studies that identified AD from one data source [2–4]. Prevalence estimates range from 2.17 million [5] to 4.5 million [6] individuals with AD in 2000, and projected prevalence estimates range from 7.98 million [5] to 13.2 million [6] individuals with AD in 2050. Depending on the populations of interest, the proportion of individuals aged 65 years and older with AD range from 0.76% [7] to 3.1% [8] among Medicare beneficiaries, from 0.83% [9] to 4.4% [10] in managed care organization populations, and from 5.7% [5] to 10.3% [11] in the general population. Given the range of prevalence estimates, costs of AD also have varied considerably. Estimates of inflation-adjusted total (direct plus indirect) cost per person from 21 studies using 1985–2000 data varied from $1,500 to $91,000 per year, equivalent to $5.6 to $88.3 billion

*Corresponding author. Tel.: 617-636-4616; Fax: 617-636-5560.
E-mail address: plin@tuftsmedicalcenter.org

1552-5260/$ – see front matter 
doi:10.1016/j.jalz.2009.09.001

nationally, primarily as a result of differences in data sources and methods for defining AD [12,13].

Although administrative data provide a more objective means than patient self-reports for identifying AD cases, potential underdiagnosis and undercoding of AD might lead to substantial underestimates of disease prevalence [13,14]. The coding bias also is reflected by the observation that the diagnosis coded in claims files might be the condition that is more likely to result in payment than AD [15,16]. In an extension of claim-based studies, Newcomer et al [15] found that fewer than 20% of a Medicare sample known to have some form of dementia were classified as having AD on the basis of their claims records, whereas 68% actually had an AD diagnosis from a referring physician at the time of entering the study. Because there is no uniformly accepted definition of AD in observational studies (ie, there is no gold standard), Pressley et al [17] examined the agreement between different case definitions by using self-report or proxy-report from the 1991–1994 National Long-Term Care Survey (ie, ''Has a doctor ever told you that you had Alzheimer's disease or dementia?''), the Short Portable Mental Status Questionnaire, dementia diagnoses in linked Medicare claims, or some combination thereof. Agreement between case definitions measured by kappa coefficient was low, in the range of 0.15 to 0.41. The authors argued that relying on a single information source to forecast national dementia-related resource use might underestimate future needs because individual data sources might miss dementia cases indicated by other sources.

Uncertainties in the burden-of-illness estimates, as a result of lacking a gold standard for defining AD, make it difficult to assess health care needs and to conduct public health planning for individuals with AD [18]. This study is an exploratory exercise that extends the existing literature by examining agreement between AD case definitions that include survey report and diagnoses in claims data and the use of AD-targeted pharmacologic agents (eg, acetylcholinesterase inhibitors, memantine) that were not available at the time of the investigations of Newcomer et al [15] or Pressley et al [17]. Using more recent data from the 1999–2004 Medicare Current Beneficiary Survey (MCBS) enabled us to provide more current estimates. We also estimated Medicare expenditures by using these different case ascertainment approaches to explore the extent to which health and drug expenditures vary by case definition. Results from this study highlight the need to improve the precision in AD case ascertainment and estimates of expenditures related to individuals with AD to support policy initiatives and quality improvement initiatives.

## 2. Methods

### 2.1. Data source

This study used the 1999–2004 Cost and Use files from the MCBS linked to corresponding Medicare Part A (ie, hospitalization, skilled nursing facility, hospice, home health care) and Part B (ie, physician visits, outpatient care) claims records [19]. The data set provides a unique opportunity to examine different definitions for selecting individuals with AD because it integrates survey information, which can be obtained directly only from a beneficiary or an appropriate proxy respondent, with reliable claims data that include diagnosis codes, utilization, charges, and reimbursement for all services rendered [20]. Pharmacy claims data were not available in 1999 to 2004. Information on AD prescription drug use was obtained from survey report and ascertained during face-to-face interviews supplemented by visual verification (ie, examination of prescription containers, pharmacy bags) of the corresponding medication to increase data accuracy.

### 2.2. Sample

The study sample consisted of community-dwelling Medicare beneficiaries aged 65 and older (n = 57,669). Nonelderly beneficiaries and facility residents were excluded (n = 11,423). We categorized eligible beneficiaries as having AD on the basis of the following three definitions:

(1) Affirmative answer to the question ''Has a doctor ever told you that you had Alzheimer's disease or dementia?'' in survey data;
(2) At least one International Classification of Diseases-ninth revision-Clinical Modification (ICD-9-CM) diagnosis code indicating AD in Medicare Part A or Part B claims files: any 290 codes (senile and presenile organic psychotic conditions) or 331.0 (AD); or
(3) Use of any AD-targeted prescription drugs, including donepezil (Aricept), rivastigmine (Exelon), galantamine (Reminyl or Razadyne), and memantine (Namenda). These medications were identified by the drug names in survey data.

Then, individuals with AD were classified into six groups: (1) AD by survey report, (2) AD by ICD-9-CM diagnosis codes in Medicare claims, (3) AD by reported use of any AD-targeted medications, (4) AD indicated by at least two definitions, (5) AD indicated by all three definitions, and (6) AD identified by any of the three definitions. Individuals with negative answers to all three case definitions were categorized as not having AD.

### 2.3. Measures

We explored discrepancies between beneficiaries captured by different case definitions by comparing their demographic characteristics (ie, age, gender, and race/ethnicity) and health-related characteristics. We also assessed discrepancies in these characteristics between individuals defined as having AD versus not having AD to provide a point of comparison. We examined functional disability, measured by six basic activities of daily living (ADLs) and six instrumental ADLs (IADLs), because it is a critical part of the disease progression and is correlated strongly with health