PREDICTIVE MODELING OF HEALTH CARE EXPENDITURES FOR MEDICARE BENEFICIARIES WITH ALZHEIMER'S DISEASE

Pei-Jung Lin

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Health Policy and Administration, School of Public Health

Chapel Hill
2008

Approved by:

Andrea K. Biddle

Rahul Ganguly

Daniel I. Kaufer

Matthew L. Maciejewski

John E. Paul

ABSTRACT

PEI-JUNG LIN: Predictive Modeling of Health Care Expenditures for Medicare Beneficiaries with Alzheimer's Disease
(Under the direction of Andrea K. Biddle)

Using the 1999-2004 Medicare Current Beneficiary Survey linked with Medicare claims, this dissertation identified the most sensitive and specific alternatives for selecting who has Alzheimer's disease (AD). This dissertation examined expenditure concentration and persistence patterns among seniors with AD, and was the first, to our knowledge, to compare the performance of different risk adjustment measures to predict overall and drug expenditures in the AD population.

The use of survey-reported AD, diagnosis in medical claims, and Alzheimer-specific prescription medicine identified different subsets of seniors with AD. 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. Using at least one of the three disease markers as our case definition, overall and drug expenditures were highly concentrated and persistent over a two-year period. Prior expenditures and comorbidities, but not functional ability, were highly predictive of the level of subsequent-year expenditures.

In single-measure, diagnosis-based risk adjustment models, the Charlson Comorbidity Index outperformed the Diagnostic Cost Group-Hierarchical Condition Category used by Medicare and the Chronic Illness and Disability System-Medicare in predicting total and drug expenditures. The frailty adjuster based on limitations of activity of daily living improved overall prediction and predictive accuracy of diagnosis-based models. Only prior drug expenditures predicted drug expenditures well. Future research is needed to evaluate

the performance of risk adjustment measures based on ambulatory pharmacy data.

In conclusion, the development of a comprehensive case definition that is reasonably sensitive and specific is crucial in observational studies to identifying individuals at different phases of AD and to assessing their health care needs. With proper organization of information, risk adjustment appears promising in predicting health expenditures even in a population with substantial disabilities. Predictive models identifying seniors with AD who are at risk for higher future expenditures can help managed care organizations to mitigate costs or change care patterns through disease management, which may delay time to nursing home care. Better care coordination, such as medication management by nurses or pharmacists, is needed to improve adherence to drug therapies for concomitant conditions with AD.

ACKNOWLEDGMENTS

I would like to thank my advisor and committee chair, Dr. Andrea Biddle, for her generous time and invaluable mentorship throughout my doctoral study. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and provided me with a tempering hand when I need focus. I am indebted to Dr. Biddle for her continuous encouragement and guidance.

I wish to thank my other committee members for their support, advice, and generosity. Dr. Rahul Ganguly encouraged me to study Alzheimer's disease and provided many practical suggestions. Dr. Daniel Kaufer's insightful comments and clinical expertise helped me enrich my ideas and put my research into perspective. I am grateful to Dr. Matthew Maciejewski for holding me to a high research standard and for guiding me to meet my goals. I am also thankful to him for providing access to the Medicare Current Beneficiary Survey for this research. Dr. John Paul continually stimulated my critical thinking. This dissertation is improved greatly for his valuable feedback.

Dr. Weng-Fong Huang and Dr. Peter Neumann have inspired me and provided wise counsel over the years. I also have benefited greatly from being around many wonderful colleagues and friends, including Scott Hsu, Shih-Yen Pan, LC Chien, and members of the Health Policy and Administration department at UNC. A special thank goes to the UNC/GlaxoSmithKline Health Outcomes Fellowship for providing financial support.

Finally, none of this would have been possible without the love and support of my family. My parents, my sister, and my brother-in-law always have been a constant source of strength, and I thank them heartily.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

AChEI	Acetycholinesterase Inhibitor
ACG	Adjusted Clinical Group
AD	Alzheimer's Disease
ADAS-cog	Alzheimer's Disease Assessment Scale-Cognitive Subscale
ADL	Activity of Daily Living
ADRD	Alzheimer's Disease and Related Dementias
CCI	Charlson Comorbidity Index
CDS	Chronic Disease Score
CDPS	Chronic Illness and Disability System
CDPSM	Chronic Illness and Disability System-Medicare
CI	Confidence Interval
CMS	Centers for Medicaid and Medicare Services
CMS-HCC	CMS Diagnostic Cost Group-Hierarchical Condition Category
DAT	Dementia of the Alzheimer Type
DCG	Diagnostic Cost Group
DCG-HCC	Diagnostic Cost Group-Hierarchical Condition Category
FDA	Food and Drug Administration
GLM	Generalized Linear Model
GRAM	Global Risk Assessment Model
HCC	Hierarchical Condition Category
HCFA	Health Care Financing Administration
HMO	Health Maintenance Organization
IADL	Instrumental Activity of Daily Living
ICD-9-CM	International Classification of Diseases, 9^{th} Revision, Clinical Modification

IHCIS	Integrated Healthcare Information Services
MCBS	Medicare Current Beneficiary Survey
MCI	Mild Cognitive Impairment
MCO	Managed Care Organizations
MEPS	Medical Expenditure Panel Survey
MMSE	Mini-Mental State Examination
NDC	National Drug Codes
NPV	Negative Predictive Value
OLS	Ordinary Least Squares
PACE	Program of All Inclusive Care for the Elderly
PHD	Pharmacy Health Dimensions
PIPDCG	Principal Inpatient Diagnostic Cost Groups
PPV	Positive Predictive Value
ROC	Receiver Operating Characteristic
SF-12	12-Item Short-Form Health Survey
SIC	Seattle Index of Comorbidity
SD	Standard Deviation
SPMSQ	Short Portable Mental Status Questionnaire
TANF	Temporary Assistance to Needy Family
VA	Veterans Affairs

CHAPTER 1

INTRODUCTION

Predictive modeling increasingly has been used to explain variation in concurrent or prospective health care utilization, especially in managed care systems. As health care spending for Medicare beneficiaries continues to rise, there is a critical need to identify high-risk populations, predict resource use, and control costs using tools such as risk adjustment models. However, observers report that Medicare risk-adjusted capitation models do not adequately compensate health plans serving primarily disabled or frail populations [Riley, 2000; Kautter and Pope, 2004]. Improvements to risk adjustment models are necessary for certain Medicare subpopulations, such as individuals with functional impairments.

This study examined the dynamics of overall health care and prescription drug expenditures for Medicare beneficiaries with Alzheimer's disease (AD), a population with progressive disabilities and high expenditures. The following sections summarize the burden of AD in the United States and describe the policy significance, followed by the purpose and structure of this dissertation.

1.1 Burden of Alzheimer's Disease

AD is a neurodegenerative disease characterized by a progressive decline in cognition, memory and functional ability. Individuals with AD often experience prominent and multiple symptoms, including psychological symptoms and behavioral disturbances, that are

distressing and place substantial financial burdens both on individuals and the health care system [Grossberg, 2002; Fillit and Hill, 2004; Kaufer et al., 2005]. AD is the most common form of dementia in the elderly, comprising approximately half of all dementia cases [Katzman, 1986]. Although currently there is no cure for AD, drug and non-drug treatments may help with both cognitive and behavioral symptoms. As of 2007, there are only four Food and Drug Administration (FDA) approved prescription medicines for AD. Mild to moderate AD is typically treated with acetycholinesterase inhibitors (AChEIs), including donepezil (Aricept®), rivastigmine (Exelon®), and galantamine (Reminyl® or Razadyne®). Moderate-to-severe cases often are treated with a N-methyl-D-aspartate receptor antagonist, memantine (Namenda®), alone or as a combination therapy with AChEIs. These drugs may temporarily delay memory decline but none of them is known to stop the underlying degeneration of brain cells [Alzheimer's Association, 2007].

In randomized controlled trials, cognitive function commonly is measured by the Mini-Mental State Examination (MMSE) scores or the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog). The MMSE evaluates cognitive decline on a scale of 0–30, with a higher score indicating less impairment [Folstein, 1975]. MMSE scores are calculated by summing subscores for seven domains: orientation to time, orientation to place, registration, recall, attention/concentration, language, and figure copy representing constructional praxis. The ADAS-cog uses 11 cognitive items, including spoken language ability, comprehension of spoken language, recall of test instructions, word-finding difficulty in spontaneous speech, following commands, naming objects and fingers, constructional praxis, ideational praxis, orientation, word-recall task and word-recognition task [Rosen et al., 1984]. The ADAS-cog scale ranges from 0 to 70, with higher scores indicating greater impairment. Individuals are categorized as having mild, moderate, or severe AD based on their cognitive function. From the time of diagnosis, individuals with AD survive about half as long as those of similar age without dementia [Larson et al., 2004]. A recent, fifteen-year prospective

epidemiological study [Ganguli et al., 2005] estimated that the mean duration of survival with AD was 5.9 (standard deviation=3.7) years; the average survival time is affected by age at diagnosis and severity of other medical conditions.

Aging is one of the greatest risk factors for AD. Approximately one in ten individuals older than 65 and nearly half of those older than 85 are affected [Evans et al., 1989]. AD has gained importance as a cause of death among older persons over the past two decades. In 2002, it was the sixth leading cause of death for persons older than 65; approximately 3.2% of all deaths among older persons were attributable to AD [Gorina et al., 2006]. The disease affects approximately 4.5 million Americans and this number is projected to increase to 11.3-16 million by 2050 [Evans, 1990; Hebert et al., 2003].

The economic burden of AD is well recognized in the literature. Bloom and colleagues reviewed 21 cost-of-illness studies on AD using 1985–2000 data and found that annual inflation-adjusted total expenditures for AD varied from $5.6 billion to $88.3 billion [Bloom et al., 2003]. Total (direct plus indirect) expenditures per patient varied from $1,500 to $91,000; indirect/family expenditures varied from $3,700 to $21,000. Although these estimates are widely variable due to differences in study design, data sources and services included, Bloom's study indicated that the cost of AD are likely to increase given the aging population. Medicare pays for most of the cost of hospitalization and a large portion of other medical care among individuals with AD. In 2000, individuals with AD represented fewer than 5% of Medicare beneficiaries [Taylor and Sloan, 2000], whereas they accounted for 14.4% of overall Medicare spending [Alzheimer's Association, 2001]. According to a report commissioned by the Alzheimer's Association, total health expenditures of Medicare beneficiaries with AD are expected to increase 75%, from $91 billion in 2005 to $160 billion in 2010; Medicaid expenditures for residential dementia care will increase 14%, from $21 billion in 2005 to $24 billion in 2010 [Alzheimer's Association, 2004].

1.2 Policy Significance

Accurate estimation of population-level prevalence is crucial to forecasting future burden of AD and to assessing the health care needs. Although prevalence estimates consistently shows an increasing trend over time and by advancing age, the figures vary widely due to data source and methodological differences [Johnson et al., 2000; Rice et al., 2001]. Using a single measure, such as an AD diagnosis in administrative data, to identify individuals with AD has been shown to introduce errors of omission and commission [Newcomer et al., 1999; Fillit, 2000; Rice et al., 2001; Pressley et al., 2003]. This dissertation investigated the utility of two supplements to claims data, Alzheimer's medication use and survey report of AD, to provide greater confidence in case ascertainment. The extent to which the costs of AD vary by case definition also was examined to provide insight into causes of wide cost variations, and to further discussions about the need for improving cost estimates to aid planning AD policy initiatives.

Additionally, despite numerous cost-of-illness studies on AD, little is known about the characteristics of individuals with AD who accrue high expenditures, and who continue to spend a disproportionately large share of money. This dissertation examined the concentration and persistence of overall and prescription drug expenditures in individuals with AD, and evaluated person-level characteristics, such as functional ability and comorbidities, to predict the level of future expenditures. Results from this study may further our knowledge of how expected high expenditures in individuals with AD may be reduced with improved care coordination and effective disease management.

Building on the foundation of accurate identification of individuals with AD and better understanding of high-expenditure concentration and persistence, this dissertation next evaluated various risk adjustment measures in predicting overall and drug expenditure. Risk adjustment models have been used to set capitation payments with the expectation that they will mitigate financial consequences of adverse risk selection into particular health plans. The

technique also is used to identify prospectively future high-risk, high-cost cases. However, a concern has been raised about whether predictive models using only diagnoses in claims data (i.e., diagnosis-based approaches) can make adequate risk-adjusted payments to Medicare plans, especially plans mainly serving frail or disabled beneficiaries. To improve model performance, researchers have suggested the inclusion of self-reported health from survey data as predictors of future expenditures [Lamers, 1999; Pacala et al., 2003]. However, more research is needed to refine risk adjustment models for certain Medicare subpopulations, such as individuals with functional limitations. Because individuals with AD generally have some level of functional impairment, they represent a distinct group for examining the performance of various risk adjustment measures (e.g., demographic characteristics, diagnoses in medical claims, and functional status from survey data). This dissertation provided recommendations for improving the predictive accuracy and usefulness of risk adjustment models in the context of predicting future expenditures. Additionally, the acquisition, cleaning, and use of medical cost data are often time-consuming and costly for health plans and employers [Powers et al., 2005]. This analysis also provides guidance to managed care organizations (MCOs) on routinely collecting selective, important risk adjustment measures at the individual level.

1.3 Purpose of This Dissertation

The purpose of this dissertation is to evaluate the following primary policy question: How well do current risk adjustment models predict overall and prescription drug expenditures for Medicare beneficiaries with AD and how can we improve the existing models? In order to answer the primary research question, this dissertation has three specific aims, with the first two aims being the foundation for the third.

Aim 1. Improve insight into the identification of community-dwelling individuals with AD in observational data

A concern has been raised about the use of a single measure, such as specific diagnosis codes in claims data, to identify community residents with AD. Potential underdiagnosis and undercoding in administrative data result in substantial underestimates of prevalence, which, in turn, leads to underestimates of the total expenditures for treating AD [Rice et al., 2001; Pressley et al., 2003]. This is of particular concern for both payers and providers who rely on claims records to determine capitation rates. Aim 1 sought to improve existing Alzheimer's case-finding methodologies by identifying the most sensitive and specific alternatives for case ascertainment, and to provide insight on causes of widely-varied expenditure estimates in the cost-of-illness literature.

Aim 2. Investigate the extent of high-expenditure concentration and persistence among individuals with AD

Aim 2 focused on an important risk adjustment measure, prior expenditures, and examined whether high-expenditure cases account for a disproportionate share of total expenditures among individuals with AD. This dissertation also sought to determine to what extent the individuals in high-expenditure percentiles persist from one year to the next and what characteristics are associated with persistence. This analysis can help to fill the knowledge gap in existing cost-of-illness studies and provide empirical rationale for developing AD-specific risk adjustment models.

Aim 3. Evaluate the performance of commonly-used risk adjustment measures in predicting expenditures among individuals with AD

Aim 3 sought to evaluate the performance of existing risk adjustment models, including diagnosis-based and survey-based measures, in predicting overall and drug

expenditures. This analysis can provide guidance on refining risk-adjusted capitation payment methods for frail Medicare beneficiaries.

1.4 Structure of This Dissertation

Chapter 2 reviews approaches to identifying individuals with AD from observational data and the limitations of existing studies. A literature review of high-expenditure concentration and persistence as well as risk adjustment modeling also is presented in Chapter 2. These sections are intended to provide background on and justification for this dissertation.

The data source and research design used in this dissertation are presented in Chapter 3. This chapter also contains a detailed description of study hypotheses and the analysis plans. Chapters 4 through 6 are individual manuscripts, each accomplishing one of the three study aims described in Section 1.3. These manuscripts take the place of a single results chapter, and each is intended for submission for independent publication. The references for the manuscripts are combined with other more general references in a single bibliography. Finally, Chapter 7 concludes this dissertation summarizing results and providing recommendations for future research and policy implications.

CHAPTER 2

LITERATURE REVIEW

This chapter describes the limitations of existing approaches to identifying community residents with Alzheimer's disease (AD) in observational data and highlights the need for improving the case-finding methodology. An overview of concentration and persistence in health care expenditures then is presented, followed by a review of risk adjustment models, applications and the associated limitations. Finally, a conceptual framework adapted from the "Algebra of Effectiveness" model [Iezzoni, 2003b] is provided to illustrate variables of interest and associations among them.

2.1 Identifying Individuals with AD from Observational Data

Validly identifying individuals with AD is a foundational step to estimate prevalence and costs of the disease. Researchers have employed several different sources of measures, such as self or proxy report, diagnosis codes in medical claims, and cognitive screening test, to define AD in observational data sets. For instance, self- or proxy-reported AD in survey data is easy to use. Individuals are assigned an AD status based on the question "Has a doctor ever told you that you had Alzheimer's disease or dementia?" Kane and Atherly used the 1991-1995 Medicare Current Beneficiary Survey (MCBS) to compare the use of Medicare covered services for individuals who reported having a diagnosis of AD or dementia with those who did not, adjusting for the prevalence of comorbidities and functional limitations [Kane and Atherly, 2000]. Based on self report, the estimated prevalence rates of AD or dementia were 1.30 for the 65–74 age group, 5.75 for 75–84, and 19.88 for 85 and older per 100 population; the overall prevalence rate was 6.8 per 100 population for women and 3.93

per 100 population for men.

Hill et al. used the 1995-1998 MCBS to investigate the relationship between the degree of functional impairment in individuals with AD or other dementia as well as their healthcare expenditures and prevalence of institutionalization in the absence of treatment [Hill et al., 2006]. Individuals with AD or dementia were defined as persons who reported having an AD or dementia diagnosis, excluding those receiving tacrine or donepezil (i.e., the two therapies available for treatment of AD during the study period). This study sample was equivalent to between 1.8 and 2 million Medicare members in each of the four calendar years in the study period.

Other researchers have relied on diagnosis codes in medical claims data to identify individuals with AD for the purpose of determining the prevalence and cost of the illness. Weiner and colleagues used claims files to analyze health care expenditures and utilization patterns for Medicare beneficiaries with dementia of the Alzheimer type (DAT), comparing with those of all Medicare beneficiaries [Weiner et al., 1998]. International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) code of 331.0 in Medicare claims was used to define the DAT sample. In fiscal year 1992, 9,323 out of 1,221,615 beneficiaries, or 0.76%, in the study had DAT; per capita Medicare expenditures for individuals with DAT were $6,208, almost twice as much as the per capita expenditure for all beneficiaries. Medicare payments increased with comorbidities, such as heart failure, chronic pulmonary diseases, and cerebrovascular disease. This study also suggests that current Medicare capitation payments to health plans may not meet the higher expected annual expenditures for caring beneficiaries with DAT.

Hill et al. carried out a retrospective analysis on administrative data for 3,934 individuals with AD and related dementias (ADRD), compared with 19,300 age/sex-matched control subjects enrolled in a large Medicare MCO [Hill et al., 2002]. Individuals with ADRD were defined based on diagnoses (ICD-9-CM codes: all 290 codes, 797, 292.82, 291.2,

294.1, 294.8, and 331.0-331.2) in medical claims and the National Drug codes (NDC) for donepezil during a two-year period. The prevalence of ADRD was 4.4%, substantially higher than reported in previous studies of Medicare MCOs. This study also analyzed the relationship between comorbidities and costs for individuals with ADRD. Their annual expenditures were found 1.6-times higher than individuals without ADRD ($10,723 vs. $6,589), after controlling for age, sex and comorbidities. Moreover, health care expenditures were even higher for individuals with ADRD who had other comorbid conditions, such as diabetes and congestive heart failure. Hill and colleagues' study suggests that better treatment and care management of ADRD may reduce the costs of managing comorbidities experienced by the frail elderly.

On the other hand, Taylor and Sloan argued that using only one or two years of claims data to identify individuals with AD may underestimate the prevalence and bias cost estimates [Taylor and Sloan, 2000]. They defined Alzheimer's cases using ICD-9-CM code 331.0 in the 1994 National Long Term Care Survey merged with a twelve-year (1984-1995) period of Medicare claims history. The effect of time since diagnosis on health care expenditures was analyzed. Approximately 3.1% of the Medicare population had a claims-based diagnosis of AD; their average total annual expenditures were 2.6-times higher compared to individuals without an AD diagnosis ($6,021 vs. $2,310 per year, $p<0.001$).

In addition, a concern has been raised about the prevalence and cost estimates when AD is defined based solely on claims records [Newcomer et al., 1999; Rice et al., 2001; Fillit et al., 2002a; Pressley et al., 2003]. Using the Medicare Alzheimer's Disease Demonstration, a longitudinal sample known to have some form of dementia, Newcomer and colleagues tested the adequacy of claims data for identifying individuals with DAT (defined by ICD-9-CM code: 331.0) and examined biases in expenditure differences between those flagged or not flagged by diagnoses using up to 36 months of claims data. Claim-based AD prevalence was underestimated substantially because of underreporting chronic conditions in claims files.

Specifically, less than 20% of the sample was classified DAT based on a single year of claims, whereas 68% actually had an AD diagnosis from a referring physician at the time of entering the demonstration program.

Many mild AD cases are underdiagnosed and, therefore, not coded in claims data [Fillit, 2000; Rice et al., 2001]. Rice et al. conducted a literature review on the prevalence, costs, and treatment of ADRD, focusing on the economics of the illness from a managed care perspective [Rice et al., 2001]. According to the US General Accounting Office (GAO) [1998], the prevalence of AD in 1995 was approximately 5.7% among those aged 65 and older. The prevalence of dementia in a Medicaid population (i.e., "Medi-Cal") was estimated at 4%, with only 1.1% having an AD diagnosis [Menzin et al., 1999]. Among the studies examined in Rice's review, the prevalence estimates from managed care claims are lower. For instance, the prevalence of diagnosed ADRD was only 0.86% among those aged 60 years and older in an MCO population [Gutterman et al., 1999]. In addition to data source and methodological differences, widely varied prevalence and expenditure estimates, in part, are attributable to the choice of diagnosis code. Different coding in medical claims may be influenced by reimbursement scheme, which provides few financial incentives for coding AD as the primary diagnosis and instead encourages coding comorbidities, such as aspiration pneumonia [Newcomer et al., 1999; Fillit, 2000; Rice et al., 2001]. Table 2.1 summarizes the AD prevalence estimates in the elderly.

Poor agreement in AD determination across different sets of measures also was observed in an analysis of a nationally representative sample of community-dwelling elderly people from the National Long-Term Care Survey data linked with Medicare claims [Pressley et al., 2003]. Dementia cases were defined by: 1) survey report; 2) moderate or severe impairment as measured on the cognitive screening test (i.e., Short Portable Mental Status Questionnaire, SPMSQ); 3) dementia diagnoses in Medicare claims (ICD-9-CM codes: 290.0-290.4, 291.2, 294.1, 331.0-331.2 and 797); 4) overlapping reports of dementia from at

least two measures; or 5) any evidence of dementia based on either survey report, SPMSQ, or diagnosis codes. The disease prevalence among individuals 65 years of age or older was found to range from 4.5% to 16.8% across the five measures; the kappa coefficient measuring agreement in case identification was low, ranging from 0.15 to 0.41. Because the non-overlapping cases (i.e., individuals identified in only one data source) also were validly determined, the authors argued that these cases should not be excluded or considered separately just because they were not explicitly identified in another measure. Considering that different measures may define different subsets of individuals with AD, the use of any single source of measure is likely to fail to capture the burden of AD at the population level.

Highly variable prevalence and cost estimates, in part, are attributable to lack of a uniformly accepted definition of AD or a gold standard diagnostic test [Bloom et al., 2003]. Researchers have suggested using supplements to claims data, such as procedure codes, interviews, medical record reviews, and physician reports, in case ascertainment [Newcomer et al., 1999]. This dissertation assessed the sensitivity and specificity of three measures, including self report, Medicare claims, and Alzheimer's medications use, to provide greater confidence in case determination (Chapter 4).

Table 2.1: Summary of AD prevalence rates in elderly population

Author (Year)	Population	Sample	Case definition	Prevalence estimate
Newcomer et al. (1999)	Medicare	DAT	ICD-9-CM: 331.0	< 20% of the sample was classified DAT, whereas 68% had an AD diagnosis
Kane and Atherly (2000)	Medicare	AD or dementia	Survey report	1.3% in age 65-74, 5.75% in 75-84, and 19.88% in age 85+
Taylor and Sloan (2000)	Medicare	AD	ICD-9-CM: 331.0	3.1%
Hill et al. (2002)	Medicare MCO	ADRD	ICD-9-CM: all 290 codes, 797, 292.82, 291.2, 294.1, 294.8, and 331.0-331.2; NDC for donepezil	4.4%
Hill et al. (2006)	Medicare	AD or dementia	Self report, excluding individuals receiving tacrine or donepezil	Equivalent to 1.8-2 million Medicare members each year
Additional studies reviewed by Rice et al. (2000)				
US General Accounting Office (1998)	General	AD	Literature review and population estimates	5.7%
Weiner et al. (1998)	Medicare	DAT	ICD-9-CM: 331.0	0.76%
Menzin et al. (1999)	Medicaid	AD or selected dementias	ICD-9-CM: 331.0, 290.0. 290.1, 290.10–290.13, 290.2, 290.20, 290.21, 290.3, 290.4, 290.40–290.43, 290.9, 294.9, 331.2, 331.89, 331.9, and 797	4%
Gutterman et al. (1999)	Medicare MCO	ADRD	ICD-9-CM: all 290 codes, 797, 292.82, 291.2, 294.1, 294.8, 331.0, 331.1, 331.2	0.83%

AD: Alzheimer's disease; ADRD: Alzheimer's disease and related dementias; DAT: dementia of the Alzheimer type; ICD-9-CM: International Classification of Diseases, 9th Revision, Clinical Modification; MCO: managed care organization; NDC: National Drug Codes.

2.2 Concentration and Persistence in Health Care Expenditures

The use of health care services in the U.S. is highly concentrated (i.e., the majority of expenditures are incurred by a relatively small group of individuals). Using the 1996 Medical Expenditure Panel Survey (MEPS), Berk and Monheit examined the distribution of health care expenditures, in terms of the share of aggregate expenditures accounted for by the top spenders, in a nationally representative sample [Berk and Monheit, 2001]. The top 1% of individuals used 27% of overall health expenditures and the top 5% accounted for more than

half of health expenditures. Such a right-skewed distribution remained stable since the 1970s. Among Medicare beneficiaries, Garber and colleagues demonstrated that the top 1% of individuals accounted for nearly 20% of overall Medicare expenditures, and that the portion of expenditures attributable to this group appeared to increase over time [Garber et al., 1997]. Moreover, individuals with high expenditures in one year were likely to have higher-than-average expenditures in other years.

In line with previous studies, managed care expenditures also were found to be highly concentrated in a small group of people. An analysis examined the model sensitivity and specificity of forecasting high-expenditure users in a random sample of enrollees from three staff/group health plans [Meenan et al., 1999]. The most expensive 1% of users represented 25% of total expenditures in this sample. Using similar approaches, another study by Meenan et al. identified high-expenditure members using data from six health maintenance organizations (HMOs) [Meenan et al., 2003]. In this sample, the most expensive 0.5% and 1% of cases represented 20% and 29%, respectively, of total expenditures.

This consistent trend of expenditure concentration has motivated researchers to investigate to what extent individuals in high expenditure percentiles persist from one year to the next. Russell and Chaudhuri evaluated the inequality of medical expenditures for several years using data collected from the Rand Health Insurance Experiment [Russell and Chaudhuri, 1992]. The experiment studied a nonelderly population covered by employer health plans for periods up to five years. In any given year, individuals in the top 5% of expenditure distribution accounted for approximately 50% of total spending. However, the expenditure persistence decreased when longer periods were examined. The proportion of total expenditures accounted for by the top 5% only represented 40% of total expenditures in the subsequent year.

In an analysis of risk adjustment models for individuals with high utilization of public mental health care, the correlation coefficients for mental health and substance abuse

expenditure were examined by type of expenditure and by year [Kapur et al., 2000]. Administrative data of severely mentally ill persons were drawn from the Los Angeles County Department of Mental Health for the fiscal years 1991 to 1994. All expenditures exhibited a relatively stable correlation over time, although the correlation was somewhat smaller for outpatient expenditures than for inpatient and total expenditures.

Persistence in the top percentiles of health care expenditures may be distributed randomly or may be correlated with personal characteristics or previous expenditures and, therefore, can be predicted. In an analysis of a longitudinal panel of the MEPS data, Monheit investigated the extent of high-expenditure persistence over a two-year period [Monheit, 2003]. Of the individuals in the top 5% of the expenditure distribution in 1996, 30% remained in this group and 45% were in the top 10% in 1997. Cancer, mental disorders, diabetes, infectious diseases, and being in the top decile (i.e., 10%-tile) of the expenditure distribution in the prior year increased the probability of expenditure persistence. However, when considering all individuals in the high-expenditure group in a given year, expenditures appeared to regress to the mean, meaning that the lower-expenditure cohorts increase and the higher-expenditure cohorts decrease their expenditures (towards the group mean) over time [Beebe, 1988; Barnett et al., 2005]. For instance, individuals who were in the top 5% of the expenditure distribution in 1996 accounted for 55.9% of aggregate expenditures but only 25% in the next year.

The high-expenditure concentration and persistence phenomenon also is observed in children. Liptak et al. analyzed data from two consecutive years of the MEPS and investigated whether high-expenditure children continue to accrue high expenditures over time [Liptak et al., 2006]. Of the total of 2,938 children, the top 10% accounted for 54% of all expenditures in 2000. Almost half of the children in the top 10% in 2000 persisted in that position in 2001, whereas 12% moved into the bottom half. Children in the top 10% in 2000 were ten times more likely their peers to be in the top 10% in the next year.

Not only overall health care expenditures, but also drug spending exhibits persistence over time, especially in the elderly population. Two studies in the early 1990s examined the persistence in drug expenditures using claims data from the Pennsylvania Pharmaceutical Assistance Contract for the Elderly program [Stuart et al., 1991; Coulson and Stuart, 1992]. Evidence suggests a strong degree of persistence in prescription drug expense by the elderly, particularly among the heaviest users. In the regression models predicting future drug expenditures, past use was found to have substantial predictive power for present use, much stronger than the effect of sociodemographic characteristics. Coulson and Stuart also demonstrated that regression to the mean may take more than four years to complete when assessing the persistence in pharmaceutical use by the elderly. In the short run, the two highest-expenditure cohorts were found to be well above the overall mean, and the three lowest cohorts were still well below the group mean.

Persistence in drug expenditures also is observed in a nationally representative sample of Medicare beneficiaries. Wrobel et al. evaluated the predictability of drug expenditures using the 1999-2000 Medicare Current Beneficiary Survey (MCBS) [Wrobel et al., 2003]. Predictors included demographic characteristics, health status, and prior drug expenditures. Demographic variables in 1999 explained only 5% of the variation in drug expenditures in 2000; adding health status measures increased the adjusted R^2 to 10%-24%. Furthermore, incorporating lagged drug expenditures (i.e. drug expenditures in 1999) raised the explanatory power to 55%.

Table 2.2: Summary of concentration and persistence in overall health care expenditures and pharmacy expenditures

Author (Year)	Data source / Study population	Results
High-expenditure concentration		
Berk and Monheit (2001)	Civilian, noninstitutionalized population from MEPS	The top 1% of spenders used 27% of health care resources; the top 5% accounted for more than half of health expenditures.
Garber et al. (1997)	Medicare enrollees	▪ The top 1% of users accounted for nearly 20% of Medicare expenditures; that the portion of expenditures attributable to this group appeared to increase over time. ▪ High-expenditure individuals in one year were likely to have higher-than-average expenditures in other years.
Meenan et al. (1999)	Managed care	The top 1% of users represented 25% of total expenditures.
Meenan et al. (2003)	Health maintenance organizations	The top 0.5% and 1% of users represented 20% and 29%, respectively, of total expenditures.
High-expenditure persistence		
Russell and Chaudhuri (1992)	Nonelderly population from the Rand Health Insurance Experiment	The top 5% of users accounted for 50% of total expenditures.
Kapur (2000)	Severely mentally ill persons with high utilization of public mental health care	Total, inpatient and outpatients expenditures all exhibited stable correlation over time.
Monheit (2003)	Civilian, noninstitutionalized population from Medical Expenditure Panel Survey	▪ Of the top 5% of users in 1996, 30% remained in this group and 45% were in the top 10% of users in 1997. ▪ Those who were in the top 5% of users in 1996 accounted for 55.9% of aggregate expenditures but only 25% in the next year.
Liptak et al. (2006)	A nationally representative sample of children from Medical Expenditure Panel Survey	▪ The top 10% of the children in 2000 accounted for 54% of all expenditures. ▪ Almost half of the children in the top 10% in 2000 persisted in that position in 2001, whereas 12% moved into the bottom half. ▪ Children in the top 10% in 2000 were ten times more likely their peers to be in the top 10% in the next year.
High-expenditure persistence in pharmacy spending		
Stuart et al. (1991); Coulson and Stuart (1992)	Pennsylvania Pharmaceutical Assistance Contract for the Elderly	▪ A strong degree of persistence in prescription drug expenditures by the elderly, particularly among the heaviest users. ▪ Past drug spending had substantial predictive power for present expenditures.
Wrobel (2003)	Medicare Current Beneficiary Survey	Incorporating prior drug expenditures substantially raised explanatory power of the model predicting drug expenditures in the subsequent year.

As summarized in Table 2.2, results from these studies generally demonstrate that a small proportion of individuals account for a disproportionately large share of health care expenditures, and that a small group exhibits persistently high expenditures. However, it is unclear whether the same pattern of expenditure concentration and persistence holds in individuals with AD given that this population has substantially more functional disabilities and greater burden of comorbidities. The uneven distribution may suggest that some groups obtain excessive care with benefits not commensurate with expenditures, whereas other groups underuse medical care. Understanding the dynamics of expenditure distribution can help decision makers plan equitable health insurance strategies, such as catastrophic care, carve-outs, reinsurance, and risk adjustment [Liptak et al., 2006]. Moreover, understanding the characteristics that predict persistence can be used as a management tool to help health plans identify individuals at risk of accruing high expenditures earlier in the process, and target members for intensive disease management and better care coordination [Russell and Chaudhuri, 1992; Ash et al., 2001].

2.3 Predictive Modeling and Risk Adjustment Measures

Predictive modeling increasingly has been used to forecast health care services use and expenditures in managed care [Pope et al., 2000; Powers et al., 2005; Zhao et al., 2005]. One of the most important applications of predictive models is for risk-adjusted payment systems. The Balanced Budget Act of 1997 required the Centers for Medicaid and Medicare Services (CMS, formerly Health Care Financing Administration [HCFA]) to implement risk-adjusted Medicare capitation payments for managed care plans by January 1, 2000, in order to fairly compensate health plans for the expected costs associated with the disease burden of their enrollees. CMS first adopted a form of risk adjustment based on inpatient hospital diagnoses (i.e., principal inpatient diagnostic cost groups system, PIPDCG) [Pope et al., 2000]. Beginning 2004, the CMS applied Medicare-specific modifications to the

Diagnostic Cost Group-Hierarchical Condition Category (DCG-HCC) model, which incorporates both inpatient and outpatient claims, as the new risk adjustment capitation payment system (hereafter known as CMS-HCC). Some states, such as Maryland, Colorado, Oregon, Delaware, Utah, Minnesota and Michigan, also use diagnosis-based classification systems to make risk-adjusted capitation payments for beneficiaries with disabilities or who receive Temporary Assistance to Needy Families (TANF) [Kronick et al., 2000].

Without risk adjustment, payers using the capitation payment methods are likely to overpay (or underpay) providers for healthier beneficiaries (or sicker groups). If payers do not reimburse more money to providers who serve enrollees with above-average levels of health care needs, plans will be penalized for attracting sicker individuals and quality of care may be jeopardized [Kronick et al., 2000]. On the other hand, without adequate payment adjustment, potential adverse selection problems may lead health plans to "enroll the healthy and spurn the sick" [Newhouse et al., 1997].

The following sections review commonly used risk adjustment models and how they are used in health care settings, with a special emphasis on elucidating the limitations of these models and the need for further refinement.

2.3.1 Risk Adjustment Models

Risk adjustment models generally use regression modeling techniques to predict outcomes (e.g., health care expenditures). The models differ from each other in the characteristics used to explain expenditures and how this information is organized in the classification system [Greenwald, 2000]. This dissertation evaluated the performance of various risk adjustment measures, including diagnoses in medical claims, functional status from survey data, and prior expenditures, in predicting future expenditures. Operational details of these risk adjustment measures are described in Section 3.7.

Studies have examined various risk adjusters, such as diagnoses [Ellis et al., 1996; Meenan et al., 1999; Ash et al., 2000; Pope et al., 2000; Riley, 2000; Meenan et al., 2003; Hughes et al., 2004; Pope et al., 2004; Noyes et al., 2006], pharmacy claims data [Von Korff et al., 1992; Fishman et al., 2003; Powers et al., 2005], prior utilization [Ash et al., 2001; Monheit, 2003], health status from survey data [Epstein and Cumella, 1988; Fowles et al., 1996; Pope et al., 1998; Lamers, 1999; Temkin-Greener et al., 2001; Pacala et al., 2003; Kautter and Pope, 2004; Fleishman et al., 2006]], or some combination thereof [Zhao et al., 2001; Maciejewski et al., 2005; Zhao et al., 2005; Farley et al., 2006]. Table 2.3 summarizes the main findings of these risk adjustment studies.

Studies focusing on diagnosis-based measures

Using 1991-1992 data for a 5% Medicare sample, Ellis and colleagues evaluated risk adjustment models that used diagnoses from both inpatient and ambulatory claims to adjust payments for aged and disabled Medicare enrollees [Ellis et al., 1996]. Hierarchical coexisting conditions (HCC) models achieved greater explanatory power than DCG models by taking account of multiple coexisting medical conditions (adjusted R^2 = 8.08 and 6.34, respectively). Although more than 90% of the variation in Medicare payments was left unexplained, the authors argued that their best-performing models may account for nearly half of the explainable variation, given that a maximum explainable portion of medical expenditure variation is estimated to be 20-25% [Newhouse, 1996].

Pope et al. described the model HCFA used in 2000, known as the PIPDCG, and assessed its performance [Pope et al., 2000]. The PIPDCG model calculates each beneficiary's relative risk factor by taking into account his/her demographic characteristics and the principal medical condition that led to an inpatient admission. Relying on the single most predictive inpatient diagnosis only, this model had an R^2 of 6.2%, approximately two-thirds of the performance of Ellis and colleagues' model [1996]. PIPDCG predicted expenditures more accurately than the model with demographic information and prior-year

expense, especially for high-cost cases. For instance, among the beneficiaries with the top 1% of prior expenditures, the PIPDCG model predicted expenditures that were 47% of actual expenditures, as opposed to only 19% by the demographic model. They concluded that the PIPDCG is a conservative model focusing on the most severely ill and expensive hospital events. The authors concluded that more refined and comprehensive models can be implemented as more experience is gained with risk adjustment.

Ash and colleagues described the structure and examined the performance of a diagnosis-based model, DCG-HCC, as developed and validated on three databases (i.e., privately insured, Medicaid, and Medicare) [Ash et al., 2000]. DCG-HCC is different from the single-condition PIPDCG model in that the hierarchies are formed to reflect chronic and serious acute manifestations of particular diseases, as well as their impact on expected expenditures. The model's explanatory power measured by R^2 ranged from 8% to 20% across the three populations. This study demonstrated that the DCG-HCC characterized the disease burden of populations and predicted future resources use, and that the model predictions can help payers to establish fair risk-adjusted payments.

Using the 1991-1994 MCBS data, Riley examined the adequacy of two diagnosis-based risk adjusters, PIPDCG and DCG-HCC, for paying health plans that disproportionately enroll frail Medicare beneficiaries, defined by institutional status and by difficulty with activities of daily living (ADLs) [Riley, 2000]. Both models were found to over-predict average expenditures for unimpaired beneficiaries and to under-predict average expenses for community-dwelling beneficiaries with functional impairments. The degree of under-prediction increased with number of ADL limitations, which were associated with Medicare expenditures, but not fully captured in demographic characteristics or diagnosis profiles. The authors concluded that functional status may be an appropriate proxy for disease severity; further refinements are needed if diagnosis-based models are used to make capitation payments to plans that disproportionately enroll frail beneficiaries.

The developers of the CMS-HCC described the model's principles, elements, organization, calibration, and performance in detail [Pope et al., 2004]. CMS-HCC was implemented in 2004 to adjust Medicare capitation payments to private health care plans for the health expenditure risk of their enrollees. The R^2 was 9.8% for the CMS-HCC model that used demographic variables plus hospital and physician claims. The authors concluded that the model improved the predictive accuracy of health care expenditures compared with previous models both in the general Medicare population and in subpopulations, such as beneficiaries entitled by disability, community vs. institutional residents, and new enrollees.

Accurate prediction for the frail elderly is a particularly important issue for MCOs whose models of care focus on the elderly in greatest need. Beginning in 2004, frailty-adjusted Medicare payments were applied to MCOs, such as Program of All Inclusive Care for the Elderly, Wisconsin Partnership Program, and Minnesota Senior Health Options, that specialize in providing care to the community-residing frail elderly. Kautter and Pope documented the development of the CMS frailty adjustment model, a Medicare payment approach that adjusts payments to MCOs according to the functional impairment of its community-residing enrollees [Kautter and Pope, 2004]. In the absence of frailty adjustment, the CMS-HCC model would under-predict Medicare expenditures by an average of $4,923, $1,531, and $809, respectively, for beneficiaries with 5-6, 3-4, and 1-2 impairments of ADLs, and over-predict by $697 for those with no ADL difficulties. The CMS frailty adjuster improved the explanatory power of the CMS-HCC model by 11%. They suggested that frailty adjustment could be applied to more Medicare MCOs in the future.

Kronick and colleagues developed the Chronic Illness and Disability Payment System (CDPS), a diagnostic classification system that used by many Medicaid programs to make health-based capitation payments for beneficiaries with disabilities or persons receiving Temporary Assistance to Needy Families (TANF) [Kronick et al., 2000].The developers of the CDPS analyzed claims data of seven state Medicaid programs and found that, for

beneficiaries with disabilities, the CDPS had a higher explanatory power than the DCG-HCC model (R^2 = 18.3% vs. 14.3%, respectively), and almost twice that of the Adjusted Clinical Group (ACG) model (R^2 = 18.3% vs. 9.8%, respectively). Overall the predicted expenditures in the subsequent year were very close to the actual expenditures (with slightly under-prediction for the plans with sick enrollees and over-prediction for the plan with healthy enrollees). The CDPS has a modified version which is more appropriate to predict expenditures for Medicare beneficiaries (hereafter known as CDPSM) [Kronick et al., 2002].

However, the CDPS does not show consistently superior performance across populations. Cumming et al. compared several currently available risk adjusters using claims data for a nationwide mix of preferred provider organizations (PPO) and HMO commercial employer group business [Cumming et al., 2002]. The CDPS model had a R^2 of 10.3% with a mean absolute prediction error of $2,299 (equal to 1.03 times the actual mean of $2,232), compared to DCGs with R^2 of 14.3% and a mean absolute prediction error of $2,187 (0.98 times actual mean).

Unlike previous studies evaluating various diagnosis-based risk adjustment models for the overall Medicare population, Noyes et al. examined the accuracy of CMS-HCC in a disease-specific context for beneficiaries with Parkinson's disease, a high-expenditure group with substantial disabilities [Noyes et al., 2006]. Using the 1992-2000 MCBS, this study demonstrated that, for the general Medicare population, the CMS-HCC model over-predicted medical expenditures for individuals without functional limitations by 15.58%. Expenditures were under-predicted as the level of disability increased, from 14.5% for individuals with 1-2 ADL limitations to 33.1% with 5-6 ADL limitations. However, such relationship was not observed for individuals with Parkinson's disease. The CMS-HCC predictions were no more than 10% off from the actual Medicare expenditures. At all ADL impairment levels, no statistically significant difference was detected between the actual cost ratios and ratios predicted by the CMS-HCC. In contrast with Kautter and Pope's conclusions, Noyes et al.

concluded that health plans covering mainly individuals with Parkinson's disease would be compensated fairly if the CMS-HCC payment model were applied.

In addition to making capitation payments, risk adjustment models can be used to identify proactively high-risk, high-expenditure plan enrollees for disease management programs. Meenan et al. compared the ability of three risk adjustment models to predict enrollees' high- and low-expenditure status in the subsequent year within a managed care population [Meenan et al., 1999]. Models examined included the Global Risk Assessment Model (GRAM) developed at the Kaiser Permanente Center for Health Research, a logistic version of GRAM, and a model based on prior expenditures. GRAM is a diagnosis-based model that incorporates demographic characteristics as well as classification of ICD-9-CM diagnoses from inpatient and outpatient encounter records. Judging by the receiver operating characteristic (ROC) curves, which illustrate the tradeoff between true positives (sensitivity) and false positives (1-specificity) across different high-expenditure thresholds, GRAM had better predictive accuracy in discriminating high- and low-expenditure individuals than did its comparators.

Another study by Meenan and colleagues evaluated the performance of five risk adjustment models, including GRAM, DCGs, ACGs, RxRisk (a model based on outpatient prescription drug fills), and prior expenditures, to identify high-expenditures individuals in an overall managed care population and within certain demographic subgroups, such as enrollees over 64, children under 13, and Medicaid recipients [Meenan et al., 2003]. They also applied the risk adjustment models to three disease groups amenable to case management: asthma, diabetes and depression. All models except RxRisk had comparably good discrimination ability (i.e., the area under the ROC: 0.83-0.86). GRAM and DCGs captured the largest proportion of total expenditures. Identifying high-risk, high-expenditure individuals is an important application of risk adjustment models. This is more useful for targeting individuals with particular chronic disease in which case management has been

shown effective, as opposed to identifying “generic” high-expenditures members. For individuals with AD who have multiple comorbidities, MCOs have many opportunities to implement interventions for improving health outcomes and lowering expenditures, including the use of appropriate medications (e.g., cholinesterase inhibitors), care coordination by nurse case managers and/or primary care physicians (e.g., discharge planning specific to AD), and education, counseling and support for caregivers [Fillit et al., 2002b; Kaufer et al., 2005].

Studies focusing on pharmacy-based measures

In addition to diagnoses, pharmacy data have been used in risk adjustment. Von Kroff and colleagues utilized patterns of use of selected prescription medications during a one year time period to construct the Chronic Disease Score (CDS) [Von Korff et al., 1992]. They concluded that scoring automated pharmacy data can provide a stable measure of chronic disease status that is associated with physician-rated disease severity and patient-rated health status, and is predictive of subsequent mortality and hospitalization rates, after controlling for age, gender and health care visits.

Fishman et al. developed and estimated the RxRisk model, a risk adjustment technique that uses automated ambulatory pharmacy data to identify chronic conditions and predict future health care costs [Fishman et al., 2003]. The RxRisk model's performance in predicting subsequent-year expenditures was compared with a demographic model and two diagnosis-based models (i.e., ACG and DCG-HCC). The DCG-HCC produced more accurate forecasts of total expenditures than its comparators, explaining 15.4% of the expenditure variation, whereas only 8.7% and 10.2% were accounted for by RxRisk and ACGs, respectively. However, all three models had similar predictions for the middle 60% of the expenditure distribution. The authors concluded that the pharmacy-based RxRisk is an alternative risk adjustment instrument to diagnostic-based models; depending on the nature of the application, pharmacy-based models may be an appropriate option for risk adjustment.

Powers and colleagues evaluated the performance of the Pharmacy Health Dimensions (PHD) model, which incorporates 51 drug categories classified based on pharmacy claims, age, gender, and drug expenditures in the baseline year to predict annual total health expenditures in follow-up year [Powers et al., 2005]. Several model specifications, including ordinary least squares (OLS) regression, log-transformed OLS regression with smearing estimator, and three two-part models using OLS regression, log-OLS regression with smearing estimator, and generalized linear modeling (GLM), respectively, were examined. They concluded that the PHD model derived solely from pharmacy claims data can be used to predict future total health expenditures. Using PHD with a simple OLS model may provide similar predictive accuracy in comparison to more advanced econometric models.

Studies focusing on prior expenditures

As with diagnosis and pharmacy data, prior expenditures are highly correlated with expenditures accrued in the next year [Garber et al., 1997; Pope et al., 1998]. Ash et al. [2001] compared prior expenditures with diagnosis-based methods for identifying a very small subgroup of a general population with high future expenditures (e.g. top 0.5%) that may be mitigated with medical management. Diagnosis-based risk models were found at least as powerful as prior expenditures for identifying future high-expenditure cases. The author concluded that combining expenditure and diagnosis data may be more powerful and more operationally useful for targeting the medical problems that may be managed to achieve better outcomes and lower costs. However, inclusion of prior expenditures in a prospective risk adjustment model to set payment might not be practical because of the perverse incentives created [Ash et al., 2000; Greenwald, 2000; Ash et al., 2001]. For instance, models that identify illness only through hospitalizations favor admissions, and models that pay more for high spenders encourage both appropriate and inappropriate, unnecessary services.

Studies focusing on survey-based health status

In addition to diagnosis information, pharmacy data, and prior expenditures, researchers have used health status from survey data as a predictor of future expenditures. Fowles et al. [1996] compared the performance of different health status measures, such as functional ability and chronic conditions, for risk-adjusted capitation methods in a sample of the group-network HMO in Minnesota. Capitation adjustment based solely on demographic characteristics was found to perform poorly, whereas both survey-reported health status and diagnoses predicted future expenditures twice as better as the demographic model. When predicting expenditures for groups rather than individuals, the demographic model worked well for overall average but tended to over-predict expenditures for healthier groups and under-predict sicker groups. In this HMO sample, the diagnosis-based ACG model performed better than survey-reported health status and across healthier and sicker groups.

Lamers evaluated the performance of the DCG capitation model using data from a health survey of Dutch sickness fund members [Lamers, 1999]. Using stepwise regression procedures, relevant survey variables that could improve the predictive accuracy of the DCG model were identified. Results from this study showed that predictive accuracy of the model was improved further by incorporating a subset of survey variables, including perceived health, having functional limitations, consultation with general practitioners, used of home nursing, number of prescription medications used, cancer, diabetes, and use of rheumatoid arthritis medicine.

Pacala et al. conducted a mailed survey to collect data, such as demographic characteristics, general health, life style, and use of health care, to predict future expenditures for community-dwelling Medicare beneficiaries [Pacala et al., 2003]. The predictive ratios (i.e., predicted-to-actual expenditures) for the low-expenditure individuals were 0.97 for the health status model method and 1.37 for the PIPDCG method; for the high-expenditure group, the ratios were 1.01 for the health status model and 1.06 for the

PIPDCG, suggesting that PIPDCG was more likely to over-predict future expenditures. The authors concluded that survey-based health status was at least as accurate as the diagnosis-based PIPDCG model to adjust capitation payments to Medicare+Choice plans. The difference in accuracy between the two methods was greatest in relatively healthy and low-expenditure beneficiaries.

Studies comparing a comprehensive array of risk adjustment measures

Many advances in risk adjustment methodology have been made over the past decade. Researchers sought to evaluate a comprehensive array of adjustment models, including diagnoses, health status, and prior expense, either alone or in combination for predicting health care expenditures. Zhao et al. used inpatient encounter data (i.e., inpatient only DCG-HCC) and outpatient pharmacy claims (i.e., RxGroups) to forecast subsequent-year health care expenditures in a privately insured, non-elderly population [Zhao et al., 2001]. Both models performed comparably overall; RxGroups performed better in terms of identifying a group of truly low-expenditure individuals, whereas the inpatient DCG-HCC model was superior in finding a small group with extremely high future expenditure. The best-performing model comprised both inpatient diagnoses and pharmacy data and explained 11.2% of variation in future expenditure, higher than either model alone, suggesting that properly combining pharmacy and inpatient diagnoses data is promising for predicting future expenditure.

An updated analysis by Zhao et al. found that models using both pharmacy and diagnosis data best predicted subsequent-year total health care expenditures (adjusted R^2 = 16.8% vs. 11.6% for RxGroup and 0.146 for DCG-HCC), and had higher predictive accuracy measured by predictive ratios (0.95-1.05) for subgroups with major medical conditions [Zhao et al., 2005]. Pharmacy-based models predicted future drug spending better than diagnosis-based models (adjusted R^2 = 48.2% vs. 24.3%), whereas diagnosis-based models predicted total expenditures (adjusted R^2 = 14.6% vs. 11.6%) and other medical expenditure

(adjusted R^2 = 11.6% vs. 7.1%) better than pharmacy-based models. This study suggested that combined drug and diagnostic data predicted total health care expenditure better than either type of data alone. Diagnosis-based models are more useful than pharmacy data for predicting overall and medical expenditures.

Using Veterans Affairs (VA) data, Maciejewski and colleagues [2005] examined the performance of various risk adjustment models, including five diagnosis-based techniques (i.e. Charlson Comorbidity Index, ACG, DCG-HCC, Chronic Illness and Disability Payment System, and RxRisk-V), two survey-reported health status measures (i.e., SF-36 and Seattle Index of Comorbidity [SIC]), and prior expenditures, for predicting outpatient, inpatient, and total expenditures. In this VA sample, diagnosis-based measures had better predictive ability than survey-reported measures. Specifically, the best-performing model, DCG-HCC, accounted for 7.2% of total expenditure variation, higher than its comparators. Prior outpatient expenditures predicted a substantial part of subsequent-year outpatient expenditures (adjusted R^2 = 42%). In line with previous findings, models with combined measures were found to have the best overall predictions, smaller over-prediction for low-expenditure groups, and smaller under-prediction for individuals with extremely high expenditures.

In a managed care population, Farley et al. [2006] compared various comorbidity measures to predict healthcare expenditures: two diagnosis-based models (i.e., Elixhauser and Charlson Comorbidity Index), one pharmacy-based measure (i.e., RxRisk-V), and several simple count measurements, including counts of prescriptions, physician visits, hospital claims, unique prescription classes, and diagnosis clusters. The Charlson comorbidity index and the Elixhauser index performed similarly (adjusted R^2 = 11.7% and 11.5%, respectively), whereas the pharmacy-based RxRisk-V (adjusted R^2 = 15.7%) outperformed both indices. An age- and gender-adjusted regression model that included a count of diagnosis groups was the best individual predictor of expenditures (adjusted R^2 =

18.1%). They concluded that simple count measures appear to predict future expenditures better than the comorbidity indices.

As summarized in Table 2.3, these studies generally support the claim that combined models (i.e., diagnoses plus other measures) explain more variation in expenditures and improve predictive power, relative to single-measure models (i.e., diagnoses alone). Combining diagnosis and expenditure data are more powerful and operationally useful in that the diagnostic information identifies the medical problems that may be amenable to disease management; thus, better outcomes and lower expenditures can be expected [Ash et al., 2001; Maciejewski et al., 2005]. However, further refinements are necessary to improve predictions for particular subgroups, such as individuals with functional impairments.

Table 2.3: Summary of risk adjustment studies

Author (Year)	Setting	Risk Adjustment Models	Conclusions
Studies focusing on diagnosis-based measures			
Ellis et al. (1996)	Medicare	DCG; DCG-HCC	DCG-HCC models achieved greater explanatory power than DCG models by taking account of multiple coexisting medical conditions.
Pope et al. (2000)	Medicare	Age and gender; PIPDCG; prior expense	PIPDCG is a conservative model focusing on the most severely ill and expensive hospital events. Further refinements are needed for improving risk adjustment.
Ash et al. (2000)	Privately-insured, Medicaid, Medicare	DCG-HCC	DCG-HCC characterized the disease burden of populations and predicted future resources use; the model predictions can help payers to establish fair risk-adjusted payments.
Riley et al. (2000)	Medicare	PIPDCG; DCG-HCC	Both PIPDCG and DCG-HCC over-predicted average expenditures for unimpaired beneficiaries and under-predicted average expenses for those with functional impairments. Further refinements, such as taking into account functional limitations, are needed if diagnosis-based models are used to make capitation payments to plans that disproportionately enroll frail beneficiaries.
Kronick et al. (2000)	Medicaid	CDPS, DCG-HCC	CDPS had a higher explanatory power than the DCG-HCC or the ACG model. Overall the CDPS-predicted expenditures in the subsequent year were very close to the actual expenditures
Cumming et al. (2002)	PPO and HMO	ACG, CDPS, DCG	CDPS had higher explanatory power and smaller mean absolute prediction error compared with DCGs.
Pope et al. (2004)	Medicare	DCG-HCC; CMS-HCC	CMS-HCC improved the predictive accuracy of health care expenditures compared with previous models both in the general Medicare population and in subpopulations.
Kautter and Pope (2004)	Medicare	CMS-HCC and frailty adjustment	In the absence of frailty adjustment, the CMS-HCC model would under-predict Medicare expenditures for beneficiaries with ADL impairments, and over-predict expenditures for those with no ADL difficulties. The CMS frailty adjuster improved the explanatory power of the CMS-HCC model by 11%.
Noyes et al. (2006)	Medicare beneficiaries with Parkinson's disease	CMS-HCC	For the general Medicare population, CMS-HCC over-predicted medical expenditures for individuals without functional limitations and under-predicted the expenditures as the level of disability increased. However, for patients with Parkinson's disease, the CMS-HCC predictions were no more than 10% off from the actual Medicare expenditures.
Meenan et al. (1999)	Managed care	GRAM; logistic GRAM; prior expenditures	GRAM had better predictive accuracy in discriminating high- and low-expenditure individuals than did its comparators.
Meenan et al. (2003)	Managed care	GRAM; DCG; ACG; RxRisk; prior expense	All models except RxRisk had comparably good discrimination ability. The high-expenditure cases correctly predicted by GRAM and DCGs

Author (Year)	Setting	Risk Adjustment Models	Conclusions
			captured the largest proportion of total expenditures.
Studies focusing on pharmacy-based measures			
Von Kroff et al. (1992)	HMO	CDS	Automated pharmacy data can provide a stable measure of chronic disease status that is associated with physician-rated disease severity and patient-rated health status, and is predictive of subsequent mortality and hospitalization rates.
Fishman et al. (2003)	HMO	Age and gender; ACG; DCG-HCC; RxRisk	The pharmacy-based RxRisk is an alternative risk adjustment instrument to diagnostic-based models; depending on the nature of the application, pharmacy-based models may be an appropriate option for risk adjustment.
Powers et al. (2005)	HMO	Age and gender; PHD; prior drug expenditures	The PHD model derived solely from pharmacy claims data can be used to predict future total health costs. Using PHD with a simple OLS model may provide similar predictive accuracy in comparison to more advanced econometric models.
Study focusing on prior expenditures			
Ash et al. (2001)	HMO	Prior expenditures; DCG	Diagnosis-based risk models are at least as powerful as prior expenditures for identifying future high-cost cases. Combining expenditure and diagnosis data may be more powerful and more operationally useful for targeting the medical problems that may be managed to achieve better outcomes and lower expenditures.
Studies focusing on survey-based health status			
Fowles et al. (1996)	HMO	Age and gender; ACG; SF-36	Risk adjustment using the diagnosis-based ACGs model performed better than survey-reported health status and across healthier and sicker groups.
Lamers et al. (1999)	Dutch sickness fund members	DCG; perceived health, functional limitations; other variables from survey data	Predictive accuracy of the DCG model was improved further by incorporating a subset of survey variables, including perceived health, having functional limitations, consultation with general practitioners, used of home nursing, number of prescription medications used, cancer, diabetes, and use of rheumatoid arthritis medicine.
Pacala et al. (2003)	Medicare	Age and gender; PIPDCG; general health; life style; use of health care	Survey-based health status was at least as accurate as the diagnosis-based PIPDCG model to adjust capitation payments to Medicare+Choice plans. The difference in accuracy between the two methods was greatest in relatively healthy and low-expenditure beneficiaries.
Fleishman et al. (2006)	Civilized, non-institutionalized population	Age and gender; SF-12	Survey-reported health status is useful in predicting medical expenditures. More research is needed to evaluate the extent to which the SF-12 adds predictive power over a comprehensive array of claim-based diagnosis data.

Author (Year)	Setting	Risk Adjustment Models	Conclusions
Studies comparing a comprehensive array of risk adjustment measures			
Zhao et al. (2001)	Privately-insured, non-elderly population	Inpatient only DCG-HCC; RxGroup	Properly combining pharmacy and inpatient diagnoses data is promising for predicting future expenditures.
Zhao et al. (2005)	Privately-insured, non-elderly population	DCG-HCC, RxGroup	Combined drug and diagnostic data predicted total health care expenditures better than either type of data alone. Diagnosis-based models are more useful than pharmacy data for predicting overall and medical expenditures.
Maciejewski et al. (2005)	Veterans Affairs primary care	Charlson; DCG-HCC; ACG; CDPS; RxRisk-V; SF-36; SIC; prior expense	Models with combined measures had the best overall predictions, reduced over-prediction of low-expenditure groups, and reduced under-prediction for individuals with extremely high expenditures.
Farley et al. (2006)	Managed care	Elixhauser; Charlson; RxRisk-V; counts of prescriptions, physician visits, hospital claims, unique prescription classes, and diagnosis clusters	Simple count measures appear to predict future expenditures better than the comorbidity indices.

ACG: Adjusted Clinical Group; CDPS: Chronic Illness and Disability System; CMS-HCC: Centers for Medicare and Medicaid Services modified version of Diagnostic Cost Group-Hierarchical Condition Category; DCG: Diagnostic Cost Group; DCG-HCC: Diagnostic Cost Group-Hierarchical Condition Category; GRAM: Global Risk Assessment Model; PHD: Pharmacy Health Dimensions; PIPDCG: Principal Inpatient Diagnostic Cost Groups; SIC: Seattle Index of Comorbidity.

2.3.2 Estimation Methods

The objective of this dissertation was to apply the globally estimated measures and to evaluate the performance of different measures among individuals with AD, rather than to recreate or to modify each risk adjustment classification system. In estimating prospective models, regression modeling techniques were applied to incorporate age and gender categories plus risk adjustment measures in year *t* to forecast expenditures in prediction year *t+1*. Although there are several methods for modeling skewed health care expenditures, many risk adjustment models have employed (weighted) ordinary least squares (OLS) estimation (Equation 2.1 [Wooldridge, 2003]) [Pope et al., 1998; Ash et al., 2000; Kronick et al., 2000; Riley, 2000; Ash et al., 2001; Fishman et al., 2003; Pope et al., 2004; Maciejewski et al., 2005; Zhao et al., 2005; Noyes et al., 2006], whereas other researchers have utilized two-part models (Equation 2.2) [Duan et al., 1983] to estimate mixed distributions that consist

of a mass of observations at zero (i.e., nonusers) and a right-skewed distribution for users [Pacala et al., 2003; Powers et al., 2005].

OLS regression: $E[Y] = X\beta$ (Equation 2.1)

Two-part model: $\Pr[Y>0|X] = \Phi(X\gamma,\upsilon)$ (Equation 2.2)

$E[Y|Y>0,X] = X\beta + E[\varepsilon|y>0,X] = X\beta$

Alternatively, a class of predictive models, generalized linear models (GLMs), has been introduced into the analysis of expenditure data [Blough et al., 1999; Manning and Mullahy, 2001]. As shown in Equation 2.3, GLM assumes a distribution for the underlying data with a scale for the linear procedure. The function g(.) is known as the link function. The GLM approach does not require any retransformation (as needed in log-transformed OLS regressions) in that the dependent variable is an expected value and is modeled on its original scale. Usually a gamma distribution and a log link are utilized for modeling expenditure data. Therefore, the expected expenditures are derived by exponentiation of the linear predictor.

GLM: $g[E(Y)] = X\beta$ (Equation 2.3)

OLS estimation is used commonly probably because it allows easy calculation of a risk profile expressed in dollar terms for each patient by summing across coefficients for each risk adjuster [Fishman et al., 2003]. Also, empirically, the predictive power of the OLS and two-part models is very close, whereas two-part models tend to be computationally burdensome and more difficult to interpret [Pope et al., 1998]. Buntin and Zaslavsky modeled health care expenditures incurred by Medicare beneficiaries and compared eight alternative methods, including OLS and GLM estimators and one- and two-part models [Buntin and Zaslavsky, 2004]. Buntin and Zaslavsky showed that the untransformed OLS regression perfectly predicted the overall mean and the means of the subgroups, such as beneficiaries with chronic conditions, with ADL limitations, and in poor health. Compared with the GLMs,

the standard OLS models performed fairly well in terms of mean squared error although less well in terms of mean absolute prediction error.

Powers and colleagues [2005] also suggested that a standard OLS model may provide similar predictive accuracy compared with more advanced econometric models. Powers' study evaluated several estimation methods in predicting prospective health care costs using a pharmacy claims-based risk index. Of the modeling approaches assessed, OLS had the lowest mean absolute prediction error and highest adjusted R^2; the log-transformed OLS and two-part log-OLS models did not predict expenditures accurately due to log-scale heteroskedasticity. The two-part GLM had lower adjusted R^2 but similar performance in other assessment measures compared with the OLS or two-part OLS estimators. This dissertation used OLS and other model specifications to estimate prospective risk adjustment models (details described in Chapters 3 and 6).

2.3.3 Limitations of Risk Adjustment Models

The generalizability of a risk adjustment model can vary across populations. Observers report that current diagnosis-based risk adjustment measures do not adequately predict the expenditures of the frail elderly, with frailty defined as functional impairment [Riley, 2000; Robinson and Karon, 2000; Temkin-Greener et al., 2001; Kautter and Pope, 2004]. There is insufficient evidence to draw conclusions about the predictive accuracy of different risk adjustment models in a disease-specific context. As risk adjustment models continue to be developed, further refinement is necessary for particular Medicare subpopulations, such as persons with functional limitations.

Cucciare and O'Donohue reviewed recent studies on risk adjustment models and found that, in general, diagnosis-based models are more accurate than demographic models, and can more accurately reflect expenditures of patients within disease categories [Cucciare and O'Donohue, 2006]. Also, diagnosis-based models are useful for identifying high-risk,

high-cost individuals. However, with regard to predictive power, risk adjustment models generally account for relatively little variation in future health expenditures. The majority of prospective risk adjustment models reviewed in this study have adjusted R^2 of 2%-16%. Moreover, some risk adjustment measures tend to under- or over-predict health care expenditures, which may lead to inappropriate risk-adjusted capitation payment to health plans. For example, Ash et al. [2000] found that the DCG-HCC model under-predicted health care expenditures for individuals with arthritis by as much as 36%. Kautter and Pope [2004] showed that, without frailty adjustment, the CMS-HCC model may underpay MCOs which specialize in providing care to the community-residing frail elderly.

CHAPTER 3

RESEARCH DESIGN AND METHODS

3.1 Overview

This dissertation examined multiple-source case ascertainment approaches, including survey report, diagnosis, and use of Alzheimer's prescription medicine, for identifying individuals with AD. Using an inclusive case definition to define AD, this dissertation then described patterns of expenditure distribution in terms of concentration and persistence. Student's t-tests and chi-square tests were performed to assess the statistical significance of differences in characteristics between the highest-spending 10%, 25%, and 50% versus the lowest-spending 50% of individuals with AD. A transition probability matrix illustrated the change in expenditure percentiles from a given year to the next year. Ordered logit regression techniques were used to model individual-level characteristics associated with future expenditures. The risk adjustment measures assessed in this study included three diagnosis-based measures from Medicare claims files, one functional status measure from survey data, and prior expenditures aggregated from both claims and survey data. Adjusted R^2, log likelihood values, predictive ratios, and the area under the receiver operating characteristic (ROC) curves were used to compare performance across models.

3.2 Research Hypotheses

This dissertation has five testable hypotheses. Hypothesis 1 corresponds to Aim 1 and explores the sensitivity and specificity of alternative definitions for case ascertainment. Hypotheses 2, 3 and 4 relate to Aim 2 and examine high-expenditure concentration and

persistence over-time in individuals with AD. Finally, Hypothesis 5 corresponds to Aim3, assessing the performance of various risk adjustment models in AD.

H1: Different disease definitions, including survey report of AD, AD diagnosis in medical claims records, and use of Alzheimer's medications, have poor agreement in case determination.

H2: A relatively small proportion of individuals with AD account for a disproportionately large share of overall health care and prescription drug expenditures.

H3: A large percentage of individuals with AD who are in the top percentiles for overall and drug expenditures in year *t* remain in the top percentiles in year *t+1*, whereas most individuals in the bottom half of the expenditure distribution in year *t* remain in the low-expenditure groups in year *t+1*.

H4: Greater burden of comorbidities and functional impairments, and higher prior expenditures in year *t* are associated with increased likelihood of being in the top percentiles for overall and prescription drug expenditures in year *t+1*.

H5: Risk adjustment models with combined measures (i.e., diagnoses plus other risk adjusters) have higher explanatory ability and better predictive accuracy compared with single-measure models (i.e., diagnoses alone) in predicting subsequent-year overall and drug expenditures.

3.3 Conceptual Framework

A broad range of risk adjustment factors can affect an individual's health expenditures. However, no risk adjustment methods can account for all relevant factors. From a practical standpoint, Iezzoni argued that one of the most important features of any risk adjustment approaches involves which measures to include and how they are represented and handled analytically [Iezzoni, 2003a]. The current research builds upon the concept of "algebra of

effectiveness," which recognizes outcomes as complex functions of individual-level attributes and other factors [Iezzoni, 2003b].

Figure 3.1 illustrates the "algebra of effectiveness," for which risk adjustment measures are integral. As shown in the diagram, outcomes are a function of intrinsic individual-level factors, treatment effectiveness, random events, and quality of care. Many diverse attributes contribute to individual-level factors, including both clinical (e.g., principal diagnosis) and nonclinical (e.g., patient attitudes and preferences) factors. This dissertation examined the relationship between a subset of individual factors (shown in shaded boxes) and resource use as an outcome.

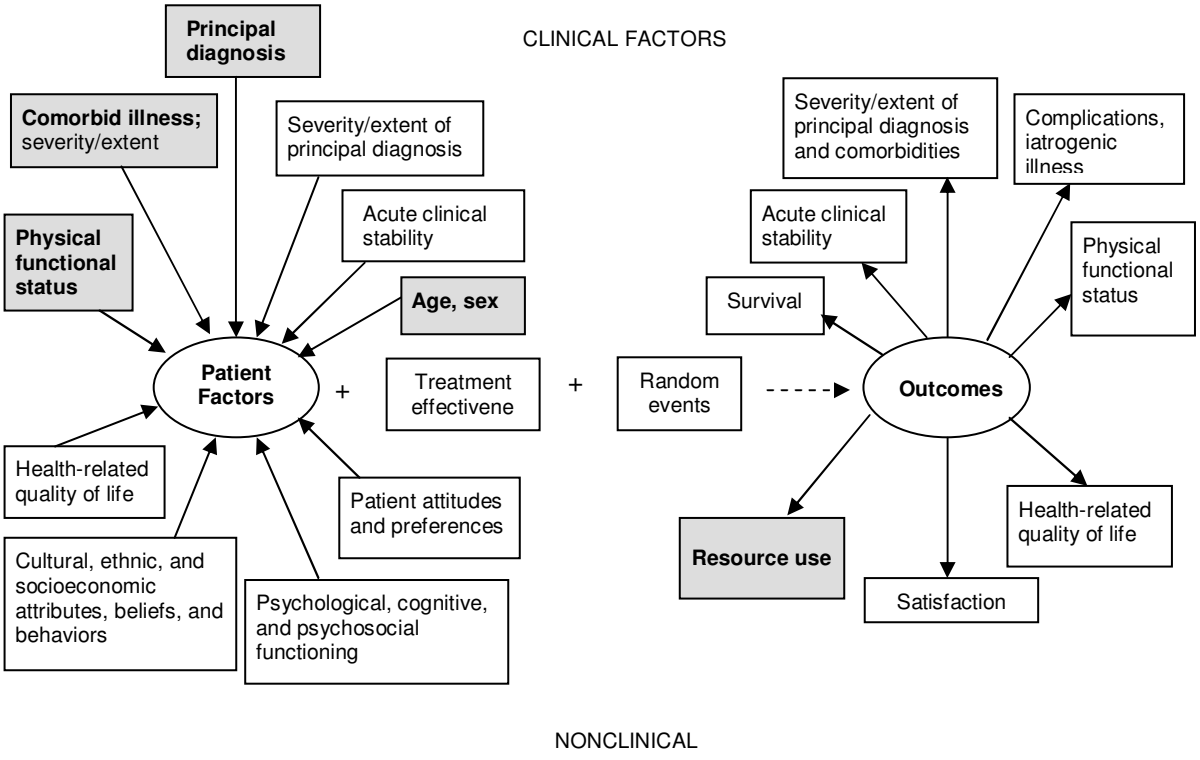


Figure 3.1: The algebra of effectiveness model
Source: Iezzoni, L. I. 2003a. "Ch.1: Reasons for risk adjustment." Risk adjustment for measuring health care outcomes, 3rd ed., Chicago, Illinois, Health Administration Press. Copyright 2003. Reprinted with permission from the Health Administration Press

Figure 3.2 shows the individual-level risk adjustment factors of interest, including age, sex, diagnoses (i.e., principal diagnosis and comorbid illness in Iezzoni's model) and survey-based functional status (i.e., physical functional status in Iezonni's model). Additionally, prior expenditures also were examined. Outcomes were measured by total and prescription drug expenditures.

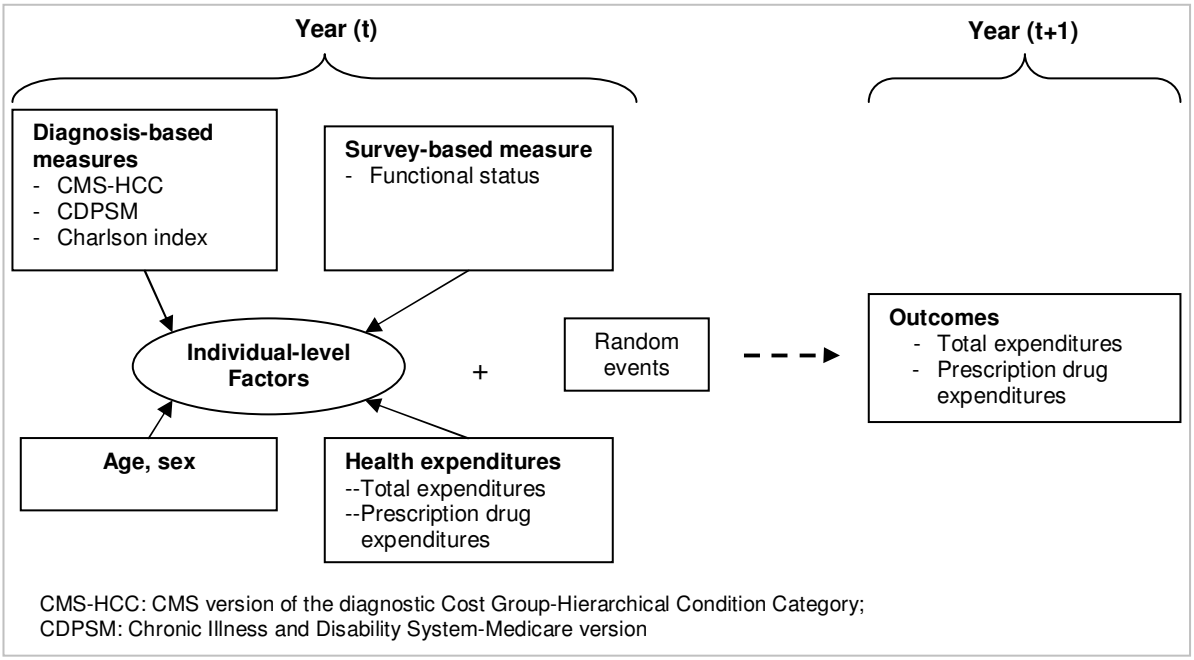


Figure 3.2: Conceptual framework

3.4 Data Source

This dissertation used data from the 1999-2004 waves of the Medicare Current Beneficiary Survey (MCBS) Cost and Use files, linked with Medicare claims data. The sample for MCBS is drawn from the Medicare enrollment file covering the entire Medicare population, whether aged or disabled, living in the community or in institutions. The data set features a nationally representative and comprehensive survey of health care use, expenditures, sources of payment, supplementary health insurance, living arrangements, income, health status, and physical functioning for the Medicare population [CMS, 2006b]. Prescription use also is obtained from survey data and ascertained during face-to-face interviews as well as

by visual verification (i.e., of prescription containers, pharmacy bags) by interviewers for the corresponding medication. Follow-ups of prescribed medicine are conducted by phone interview. Available Medicare Part A (i.e., hospitalization, skilled nursing facility, hospice, and home health care) and Part B (i.e., physician visits, specified outpatient care and some outpatient medications) claims records include diagnosis codes, utilization, charges, and reimbursement for all medical services [CMS, 2003].

Survey interviews are completed by either sample members or proxy respondents (usually a family member or close acquaintance) if the sample member is unable to respond due to physical or mental problems. Each individual is interviewed three times per year to form a continuous profile of the individual's health care experience [CMS, 2006b]. The survey features a longitudinal rotating panel design, in which sampled individuals remain in the panel for no more than four years by the time they are retired, resulting in a cumulative sample size of three 4,000-beneficiary cohorts or approximately 12,000 beneficiaries in any given year. Therefore, the MCBS Cost and Use data set can support both cross-sectional (e.g., concentration of health expenditures) and longitudinal analyses (e.g., persistence of health expenditures). Moreover, it allows us to construct risk adjustment measures using base-year information to predict subsequent-year expenditures.

3.5 Sample

The sample extraction flowchart is shown in Figure 3. Our study sample consisted of elderly, community-dwelling Medicare beneficiaries, defined as adults aged 65 and older who were not institutionalized for more than 90 days at a time during a year [CMS, 2006c] (n=69,092). Eligible beneficiaries were categorized as having AD based on the following three definitions:

1) affirmative answer to the question "Has a doctor ever told you that you had Alzheimer's disease or dementia?"; or

2) at least one AD diagnosis, defined by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes: all 290 codes (senile and presenile organic psychotic conditions) and 331.0 (AD) in Medicare Part A or Part B claims files; or

3) use of any Alzheimer's medications, including donepezil (Aricept®), rivastigmine (Exelon®), galantamine (Reminyl® or Razadyne®), and memantine (Namenda®). These medications were identified by the drug names in survey-reported data, i.e., pharmacy administration data were not used

For the case-finding analysis presented in Chapter 4, individuals with AD were classified into six categories: three single-definition groups (AD by survey report, AD by ICD-9-CM diagnosis codes in Medicare claims, and AD by any prescription mention of Alzheimer's medications), one group of AD cases indicated by at least two definitions, one group with all three definitions, and one composite group representing individuals identified by any of the three definitions. In Chapters 5 and 6, individuals who met any of the three definitions were defined as having AD. Individuals with AD who had any managed care participation during that year were excluded because these Medicare MCOs do not submit claims with diagnoses to the CMS, which were required to construct diagnosis-based comorbidity profiles. Thus, a total of 2,779 observations from 1,861 unique individuals with AD from the 1999-2004 waves of the MCBS were pooled for the expenditure concentration analysis (Chapter 5). For the expenditure persistence analysis (Chapter 5) and the risk adjustment analysis (Chapter 6), we excluded persons with only one year of claims. We then retained individuals who had medical claims data for both year *t* and year *t+1*, and excluded data for year *t+2* among those who were observed more than twice in the data set. Therefore, two-year panels of 671 unique individuals with AD were analyzed.

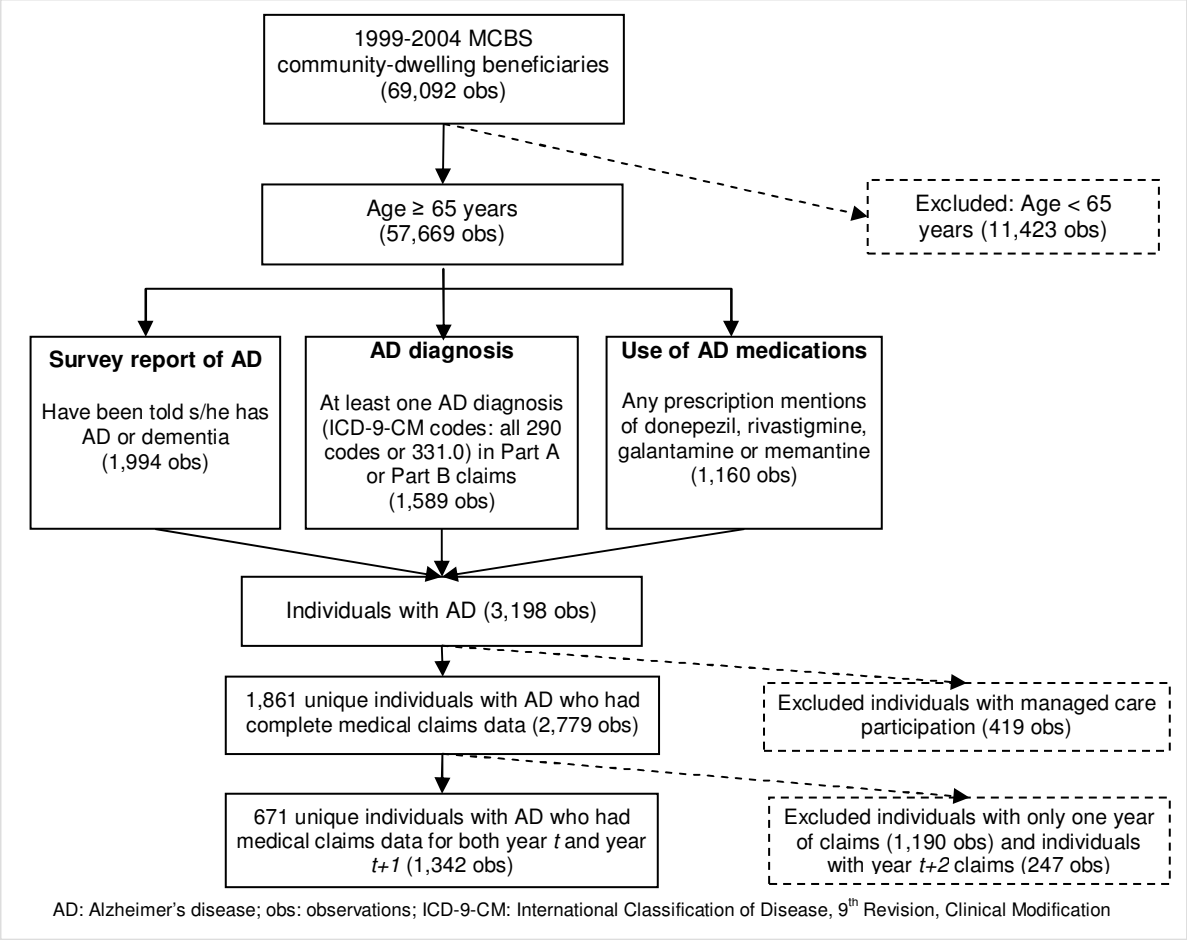


Figure 3.3: Sample extraction flow diagram

3.6 Expenditure Measures

Expenditure data in the MCBS were developed through a reconciliation process combining Medicare administrative files and survey information [CMS, 2003]. Expenditure data from claims include complete billing records of services used and payments made under Medicare fee-for-service transactions, and therefore allow for matching MCBS reports to billing records [CMS, 2006b]. The MCBS Cost and Use files link Medicare claims to survey-reported events, and thus provide a comprehensive picture of expenditures and sources of payment for all health care services, including those not covered by Medicare such as outpatient prescription drugs. Personal total health care expenditures in our study are defined as aggregated payments across all types of services, including inpatient

hospitalizations, outpatient hospital care, physician services, home health services, durable medical equipment, skilled nursing home services, hospice services, and other medical services. The payments do not include rebates that may be paid by the manufacturer to the insurer. Prescription drug expenditures are imputed total payments in the MCBS for prescription medicines received from all sources, including Medicare, Medicaid, Medicare MCOs, private MCOs, the Veterans Administration, employer-sponsored private insurance, individually-purchased private insurance, private insurance from an unknown source, out-of-pocket payments and public health plans other than Medicare or Medicaid [CMS, 2003]. Detailed procedures and criteria used in the MCBS to impute missing payments for medical services and prescription drugs are discussed elsewhere [England et al., 1994]. Briefly, a full set of internally consistent expenditure and payment records were created from these sources with very little partial data discarded. All expenditures were converted into constant 2007 dollars using the Consumer Price Index for medical care [Bureau of Labor Statistics, 2007], which reflects the average price paid for a typical market basket of health care services.

3.7 Comorbidity and Functional Status Measures

This dissertation evaluated the performance of six risk adjustment techniques in predicting future overall and drug expenditures. The simplest technique considers only age categories (e.g., 65-69, 70-74, 75-79, 80-84, and 85+) and gender; three techniques are based on diagnoses in Medicare claims data, and one is based on survey-reported functional status. The final one is prior-year expenditures. These latter five risk adjustment measures are discussed individually in Sections 3.7.1-3.7.5.

3.7.1 CMS Diagnostic Cost Group-Hierarchical Condition Category (DCG-HCC)

The DCG-HCC model uses age, sex and ICD-9-CM diagnosis codes from inpatient admissions and outpatient services to predict health care expenditures associated with various comorbidities [Pope et al., 2000]. Beginning 2000, the CMS has adopted the DCG-HCC model for Medicare risk adjustment largely because of its transparency, ease of modification, and good clinical coherence [Pope et al., 2004]. In 2004, the CMS implemented Medicare-specific modifications to the DCG-HCC model to adjust Medicare capitation payments to private health care plans (hereafter known as CMS-HCC). Under this model, individuals are assigned to multiple Hierarchical Condition Categories (HCCs) based on demographic characteristics and diagnoses in both inpatient and outpatient claims in the prior year. The model also incorporates the original reason for Medicare entitlement (disability or age), location of residence (community or facility), recent enrollment (covered by Medicare for fewer than 12 months in the prior year), and Medicaid eligibility. The model aggregates clinically and economically meaningful disease categories into 70 HCCs (i.e., metastatic cancer and acute leukemia). We employed the count of HCCs per beneficiary as a risk adjustment measure. The CMS-HCC algorithm was obtained from the CMS website [CMS, 2006a].

3.7.2 Chronic Illness and Disability Payment System-Medicare version (CDPSM)

The Chronic Illness and Disability Payment System (CDPS) is an expansion of a prior model, the Disability Payment System, developed by Kronick and colleagues [Kronick et al., 1996]. At least eight state Medicaid programs use this diagnostic classification system to make risk-adjusted capitation payments for beneficiaries with disabilities or who receive Temporary Assistance to Needy Families [Kronick et al., 2000]. We used the CDPS Medicare version (CDPSM), a modified model more appropriate to predict expenditures for Medicare beneficiaries [Kronick et al., 2002] (Table 3.1). Compared with the CMS-HCC, the CDPSM is

more conservative in counting diagnoses and, therefore, has fewer categories. The CDPSM uses ICD-9-CM codes in both inpatient and outpatient claims to create 16 major disease categories, which correspond to body systems or type of diagnosis. The major categories are divided further into 66 subcategories to reflect the level of increased expenditures associated with the diagnoses. The CDPSM software was acquired under license agreement at no charge (http://cdps.ucsd.edu, last accessed on 25 January, 2008).

Table 3.1: Chronic Illness and Disability Payment System, Medicare version

Cardiovascular	*Skin*
Very high	High
Ischemic heart disease, high	Low
Ischemic heart disease, low	
Valvular, conductive and other heart disease, medium	*Renal*
Valvular, conductive and other heart disease, low	Extra high
Valvular, conductive and other heart disease, very low	Very high
Peripheral vascular, medium	Medium
	Low
Psychiatric	Very low
High	
Medium	*Substance abuse*
Low	Low
	Very low
Skeletal and connective	
Medium	*Cancer*
Very Low	Very high
Extra Low	High
	Medium
Nervous system	Low
High	Very low
Peripheral, high	
Peripheral, low	*Metabolic*
Multiple sclerosis, muscular dystrophy and others	High
Parkinson's disease	Low
Convulsions and epilepsy	
Super low	*Cerebrovascular*
	High
Delirum and dementia	Medium
Deliruim	Low
Dementia	Very low
	Extra low
Pulmonary	
High	*Infectious disease*
Medium	Acquired immunodeficiency syndrome (AIDS)
Pneumonia, high	High
Pneumonia, low	Human immunodeficiency virus (HIV)
Chronic obstructive pulmonary disease, high	Medium
Gastrointestinal	*Hematological*
High	Very high
Ostomy	High
Medium	Medium
Low	Low
	Anemia
Diabetes	
Type 1 or 2 with rare complications	
Type 1 with common complications	
Type 1	
Type 2 with common complications	
Type 2	

3.7.3 Charlson Comorbidity Index (CCI)

The original CCI is a general comorbidity index consisting of 19 disease categories developed to predict one-year mortality using hospital chart review data [Charlson et al., 1987]. Other researchers have adapted the CCI for use with comorbidity data from

administrative databases that include ICD-9-CM diagnosis codes [Deyo et al., 1992; D'Hoore et al., 1993; Romano et al., 1993; Ghali et al., 1996; Quan et al., 2002]. We used the CCI with the Deyo modification containing 17 comorbidity categories [Charlson et al., 1987; Deyo et al., 1992]. Each condition is assigned a weight as 1, 2, 3, or 6, reflecting the magnitude of the adjusted relative risks associated with each comorbidity. The CCI score then sums the weights for all conditions, with higher numbers representing a greater burden of comorbidity. The CCI for this study was derived using the presence of various ICD-9-CM codes in Medicare claims (Table 3.2).

Table 3.2: Charlson comorbidity categories and corresponding ICD-9-CM codes

Comorbidity categories	ICD-9-CM codes	Comorbidity categories	ICD-9-CM codes
Myocardial infarction	410-410.9, 412	Diabetes	250-250.3, 250.7
Congestive heart failure	428-428.9	Diabetes with organ damage	250.4-250.6
Peripheral vascular disease	443.9, 441-441.9, 785.4, V43.4, 38.48(P)	Hemiplegia or paraplegia	344.1, 342-342.9
Cerebrovascular disease	430-438	Renal disease	582-582.9, 583-583.7, 585, 586, 588-588.9
Dementia	290-290.9, 331-331.2	All tumors including leukemia/lymphoma	140-172.9, 174-195.8, 200-208.9
Chronic obstructive pulmonary disease	490-496, 500-505, 506.4	Moderate to severe liver disease	572.2-572.8, 456.0-456.21
Rheumatologic disease	710.0, 710.1, 710.4, 714.0-714.2, 714.81, 725	Metastatic solid tumor	196-199.1
Peptic ulcer disease	531-534.9, 531.4-531.7, 532.4-532.7, 533.4-533.7, 534.4-534.7	AIDS	042-044.9
Mild liver disease	571.2, 571.4-571.49, 571.5, 571.6		

ICD-9-CM: International Classification of Diseases, 9th Revision, Clinical Modification

3.7.4 Survey-reported Functional Status

Kautter and Pope [2004] developed a CMS frailty adjustment model that incorporates functional status in addition to the CMS-HCC to adjust for capitation payments to certain health plans specializing in providing care to the community-dwelling, frail elderly. The use of functional ability frailty adjuster, in particular measured by ADLs, is preferred to other health status measures because it has good face validity and has been shown to explain Medicare expenditures not accounted for by diagnosis-based measures [Pope et al., 1998; Riley, 2000;

Kautter and Pope, 2004]. The frailty adjuster uses a scale based on the count of ADL impairments, including bathing, dressing, eating, transferring in and out of chairs, walking, and toileting, and categorized as none, 1-2, 3-4, and 5-6. For each activity, individuals were asked whether they had any difficulty performing the activity, received help with the activity, needed supervision with the activity, or were unable to perform the activity because of health problems. An individual was coded as impaired for an ADL if he or she needed any assistance doing (i.e., received help or needed supervision) or could not perform the activity.

3.7.5 Prior Expenditures

Expenditures incurred in the prior year are highly correlated with subsequent-year expenditures [Garber et al., 1997; Pope et al., 1998]. We modeled baseline expenditures as continuous variables on their original scale, and used the generalized linear model (GLM, described in the next section) as an alternative to transforming the non-normally distributed data.

3.8 Analysis Plans

Analytical strategies used in this dissertation varied by research hypotheses. For Hypothesis 1 (Chapter 4), we quantified the agreement between case definitions using raw agreement (%) illustrated by a Venn diagram and chance-corrected agreement measured by the kappa coefficient [Cohen, 1960]. McNemar's test was used to test the significance of discordance between two case definitions [McNemar, 1947]. The sensitivity (i.e., the rate of true positives) is the ability of the case definition to correctly identify AD cases, whereas the specificity (i.e., the rate of true negatives) is the ability to correctly identify those who do not have AD. Positive predictive value (PPV) represents the proportion of individuals with AD among the identified cases, whereas negative predictive value (NPV) represents the proportion of individuals without AD among those identified as not having AD [Szklo and

Nieto, 2004]. These statistics provide a measure of the validity of the three case definitions, using one of them at a time as the "gold standard" definition for AD. Student's t-tests and chi-square tests were performed to assess the statistical significance of differences in characteristics among individuals in each of the six AD groups versus non-AD individuals.

For Hypotheses 2 and 3 (Chapter 5), the study sample was rank-ordered by overall expenditures and by drug expenditures, and then categorized into four groups based on their percentile position in the expenditure distribution of a given year: 1) top 10%, 2) next 11%-25%, 3) next 26%-50%, and 4) bottom 50%. We first used cross-sectional data to examine expenditure concentration by calculating the proportion of all expenditures incurred by the top 10%, top 25% and top 50% of individuals in each year. Characteristics of individuals with AD in various high-expenditure groups (i.e., top 10%, top 25%, top 50% and bottom 50%) were examined. Student's t-tests and chi-square tests were performed to assess the statistical significance of differences in characteristics between the top 10%, top 25% and top 50% versus the lower 50% of individuals. A transition probability matrix was presented to illustrate the change in expenditure percentiles from one year to a subsequent year [Monheit, 2003; Liptak et al., 2006]. Ordered logit models [Maddala, 1983] were performed to predict the probability of being in a certain expenditure group in the next year, accounting for the ordinal nature of the outcome variable (i.e., top 10%, next 11%-25%, next 26%-50% and bottom 50%).

A comparative evaluation of various risk adjustment measures to predict overall health care and prescription drug expenditures is presented in Chapter 6. A Pearson correlation coefficient matrix was used to evaluate the correlation among various risk adjustment measures. Ordinary least squares (OLS) regressions were employed to predict expenditures in year *t+1* using each risk adjustment measure plus age-gender categories in year *t*. Because the expenditure data were non-normally distributed, we also performed GLM with a gamma variance and a log link function, which models expenditures in their natural

scale rather than in a transformed scale (e.g., log transformation) [Shwartz and Ash, 2003]. The full models are described in Equations 3.1 and 3.2.

OLS: $E[Expenditure_{t+1}] = \beta_1 Age_{it} + \beta_2 Sex_{it} + \beta_3 Diagnosis_{it} + \beta_4 ADL_{it} + \beta_5 Expenditure_{it}$

(Equation 3.1)

GLM: $g[E(Expenditure_{t+1})] = \beta_1 Age_{it} + \beta_2 Sex_{it} + \beta_3 Diagnosis_{it} + \beta_4 ADL_{it} + \beta_5 Expenditure_{it}$

(Equation 6.2)

Risk adjustment models in the all-disease context conventionally have used a split-sample design to evaluate predictive accuracy, in which a small randomly selected sample of the total study population is withheld for model validation [Ellis et al., 1996; Fishman et al., 2003; Meenan et al., 2003; Pacala et al., 2003; Powers et al., 2005]. This technique helps to avoid over-fitting to a specific sample, but requires a large sample size to make statistically meaningful comparisons between a "training sample" and a model "validation sample." Since this study focuses on an AD-specific population with a sample size of 671 unique individuals, which is not sufficiently large, alternative statistical evaluation criteria were adopted. We used adjusted R^2 from OLS models to demonstrate the proportion of total variance in the dependent variable (e.g., expenditures) accounted for by the risk adjustment model [Ellis et al., 1996]. For GLMs, we compared the log likelihood values across models, with higher numbers indicating better model fit.

Although adjusted R^2 and log likelihood values provide a summary measure of overall prediction, these statistics give little information about how well a model discriminates between high- and low-expenditure cases [Shwartz and Ash, 2003]. Therefore, an individual's actual expenditures in the prediction year (i.e., year t+1) were categorized into quartiles; then, predictive ratios (i.e., predicted expenditures divided by actual expenditures) were calculated within each quartile [Ash et al., 2000; Cucciare and O'Donohue, 2006]. If the model performs well for a population, its predictive ratio is close to one. In a prospective payment system, a predictive ratio close to one indicates that aggregate payments under the

risk adjustment model are equivalent to payments under the fee-for-service [Ellis et al., 1996].

Predictive model performance was examined further using receiver operating characteristic (ROC) curves, representing how well the model classifies individuals by illustrating the tradeoff between true positive (sensitivity) and false positive (1-specificity) across different cutpoints for various high-expenditure threshold [Meenan et al., 1999]. This is done by rank-ordering actual and predicted expenditures from high to low, and setting pre-determined percentage thresholds within each expenditure distribution to define “true” high-expenditure cases (i.e., top 10% and top 20%) [Meenan et al., 1999; Meenan et al., 2003; Weiner, 2003]. A *c*-statistic representing discrimination power then can be calculated using a nonparametric trapezoidal method to approximate the area under the ROC curve [Shwartz and Ash, 2003]. It takes values from 0 to 1 with higher numbers indicating a better model fit, except for a value of 0.5 indicating no ability to discriminate.

CHAPTER 4

IDENTIFYING COMMUNITY-DWELLING INDIVIDUALS WITH ALZHEIMER'S DISEASE FROM OBSERVATIONAL DATA

4.1 Abstract

Introduction: Using diagnosis codes in administrative data to estimate the prevalence of Alzheimer's disease (AD) has been shown to introduce errors of omission and commission. The objectives of this study were to identify the most sensitive and specific alternatives for case ascertainment and to provide insight on causes of widely-varied cost estimates in the cost-of-illness literature.

Methods: Retrospective cross-sectional analyses were performed on a nationally representative sample of elderly, community-dwelling Medicare beneficiaries using the 1999-2004 waves of the Medicare Current Beneficiary Survey (MCBS) (n=57,669). Individuals with AD were identified based on six definitions according to (1) survey report only, (2) diagnosis only, or (3) use of Alzheimer's prescription medicine only, or some combination of the three. Agreement between different case definitions was assessed using a Venn diagram, kappa coefficient, sensitivity and specificity. We also explored the extent to which estimates in overall health expenditures and drug expenses vary by how AD is defined.

Results: Using any of the three case definitions, 3,198 individuals (5.55%) were identified as having AD. 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%). Kappa coefficients were low, ranging from 0.37 to 0.40. Using diagnosis in medical claims as a definitive measure for AD (i.e., "gold standard"), survey report was specific (97.7%) but had low sensitivity (44.2%) and Alzheimer's medication use was slightly more specific (98.9%)

but less sensitive (34.8%). 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. **Conclusion:** Different case definitions identify different subsets of individuals with AD. Survey-report appears to be a highly specific approach to identifying AD without the loss of sensitivity found with Alzheimer's medication use. As a consequence of different case definitions employed, health expenditures among individuals with AD varied widely. One should exercise caution in interpreting current cost-of-illness studies and in applying these estimates to policy initiatives. A crucial first step to assessing the health care needs for individuals with AD is to develop a comprehensive definition for case ascertainment.

Keywords: Alzheimer's disease, case identification/ascertainment, Medicare, prevalence, cost, functional status

4.2 Introduction

Alzheimer's disease (AD) is the most common form of dementia in the elderly, comprising approximately half of all dementia cases [Katzman, 1986]. Albeit numerous cost-of-illness studies exist for AD, great variation exists in current and projected prevalence estimates. A meta-analysis conducted by the US General Accounting Office (GAO) estimated that 2.17 million individuals had AD in 2000 [US GAO, 1998]. This number is projected to increase to 7.98 million in 2050. Much higher estimates were reported by Evans et al. [1990] and Hebert et al. [2003]. Using data from a cohort of 32,000 individuals in East Boston, an estimated 2.88 million persons aged 65 or older had AD in 1980 [Evans, 1990]. Applying U.S. Census Bureau projections of the population growth yielded 10.3 million individuals with AD in 2050. Using a similar approach, Hebert and colleagues calculated the national prevalence of AD based on data from 3,838 residents age 65 and over in a biracial Chicago community [Hebert et al., 2003]. Their results suggest that 4.5 million persons had

AD in 2000 and that 13.2 million individuals will be affected by this disease by 2050, projections even higher than those of Evans et al. Other estimates are mid-range between the GAO's and Hebert's estimates [Brookmeyer et al., 1998; Sloane et al., 2002]. Depending upon the populations of interest, the proportion of persons age 65 years and older with AD and related dementia range from 0.76% [Weiner et al., 1998] to 3.1% [Taylor and Sloan, 2000] among Medicare beneficiaries, from 0.83% [Gutterman et al., 1999] to 4.4% [Hill et al., 2002] in managed care organization (MCO) populations, compared with 5.7% [US GAO, 1998] to 10.3% [Evans et al., 1989] in the general population.

Similar to widely-varied prevalence figures, the estimated costs of AD also vary considerably. Bloom and colleagues reviewed 21 cost-of-illness studies using 1985-2000 data and found that inflation-adjusted total (direct plus indirect) costs per patient varied from $1,500 to $91,000 per year, equivalent to $5.6 to $88.3 billion nationally [Bloom et al., 2003]. This review calls into question the usefulness of extant estimates of health care expenditures for AD. Particularly, it may be difficult to assess health care needs and to conduct public health planning for persons with AD based on these widely-varied estimates.

These widely divergent prevalence and cost estimates primarily arise from the data sources employed and methodological differences in defining Alzheimer's cases [Johnson et al., 2000; Rice et al., 2001]. Although administrative data commonly are used to conduct burden-of-illness research, some researchers question the accuracy and adequacy of AD diagnosis codes in medical claims for identifying individuals with AD [Newcomer et al., 1999; Rice et al., 2001; Fillit et al., 2002a]. Pressley and colleagues [2003] analyzed a sample of community-dwelling elderly persons from the 1991-1994 National Long-Term Care Survey data linked with Medicare claims. Dementia cases were selected using survey report (i.e., "Has a doctor ever told you that you had Alzheimer's disease or dementia?"), the Short Portable Mental Status Questionnaire, dementia diagnoses in Medicare claims (defined by the International Classification of Diseases, 9th Revision, Clinical Modification [ICD-9-CM]

codes: 290.0-290.4, 291.2, 294.1, 331.0-331.2 and 797) or some combination thereof. They found that relying on a single case definition to forecast national dementia-related resource use could produce serious underestimates of future need. Because the non-overlapping individuals also were validly determined, the authors argued that these cases should not be excluded or be considered separately just because they were not explicitly identified in another sample.

During the time frame of Pressley's study, there were few pharmacologic treatment options available for dementia. However, the availability of new pharmacologic agents, such as acetylcholinesterase inhibitors and memantine, may affect current and projected estimates of the burden of AD [Brookmeyer et al., 1998; Fillit and Hill, 2004]. Using more recent data on a nationally-representative sample of Medicare enrollees enables us to investigate the utility of two supplements to claims data, Alzheimer's medication use and survey report of AD, to provide greater confidence in case ascertainment. We also explore the extent to which overall and drug expenditures vary by case definition to provide insight into causes of wide expenditure variations, and to further discussions about the need for improving utilization estimates to aid planning AD policy initiatives.

4.3 Methods

Data Source

This study used the 1999-2004 Cost and Use files from the Medicare Current Beneficiary Survey (MCBS), which is sponsored by the Centers for Medicare and Medicaid Services (CMS). The MCBS is a comprehensive survey of health care use, expenditures, sources of payment and health status [CMS, 2006b].The survey data are linked to corresponding Medicare claims records to form the Cost and Use files. The survey features a longitudinal, rotating panel design. Each individual, or an appropriate proxy respondent (usually a family member or close acquaintance), is interviewed three times a year during a

4-year period to form a continuous profile of the subject's health care experience. Each year, one round is retired and another round is initiated, giving the survey a cumulative sample size of three 4,000-beneficiary cohorts. In any given year, the MCBS is composed of approximately 12,000 individuals selected to be representative of the entire Medicare population.

The survey-reported data include information on the cost and use of all types of medical services rendered. Prescription use also is obtained from survey report and ascertained during face-to-face interviews as well as by visual verification (i.e., of prescription containers, pharmacy bags) by interviewers for the corresponding medication. Available Medicare Part A (i.e., hospitalization, skilled nursing facility, hospice, and home health care) and Part B (i.e., physician visits, specified outpatient care and some outpatient medications) claims records include diagnosis codes, utilizations and charges and reimbursement for all medical services. The richness of the MCBS Cost and Use data provides a unique opportunity to examine various definitions for selecting individuals with AD because it integrates survey information, which can be obtained only directly from a beneficiary, with reliable information on services used and Medicare payments made from administrative billing files [CMS, 2003].

Sample

The study sample consisted of elderly, community-dwelling beneficiaries, defined as adults aged 65 and older who were not institutionalized for more than 90 days at a time during a year [CMS, 2006c], in the 1999-2004 MCBS Cost and Use files (n=69,092) (Figure 4.1). For community residents, survey interviews were completed by either sample members or proxy respondents if the sample member was unable to respond due to physical or mental problems. Eligible beneficiaries were categorized as having AD based on the following three definitions:

1) affirmative answer to the question “Has a doctor ever told you that you had Alzheimer’s disease or dementia?”; or

2) at least one AD diagnosis, defined by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes: all 290 codes (senile and presenile organic psychotic conditions) and 331.0 (AD) in Medicare Part A or Part B claims files; or

3) use of any Alzheimer’s medications, including donepezil (Aricept®), rivastigmine (Exelon®), galantamine (Reminyl® or Razadyne®) and memantine (Namenda®). These medications were identified by the drug names in survey-reported data, i.e., pharmacy administration data were not used.

Then, individuals with AD were classified into six categories: three single-definition groups (AD by survey report, AD by ICD-9-CM diagnosis codes in Medicare claims, and AD by any prescription mention of Alzheimer’s medications), one group of AD cases indicated by at least two definitions, one group meeting all three definitions, and one composite group representing individuals identified by any of the three definitions. Proportions of individuals with AD were estimated for each of these six groups. Individuals with negative answers to all three case definitions were categorized as “non-AD” (i.e., not having AD).

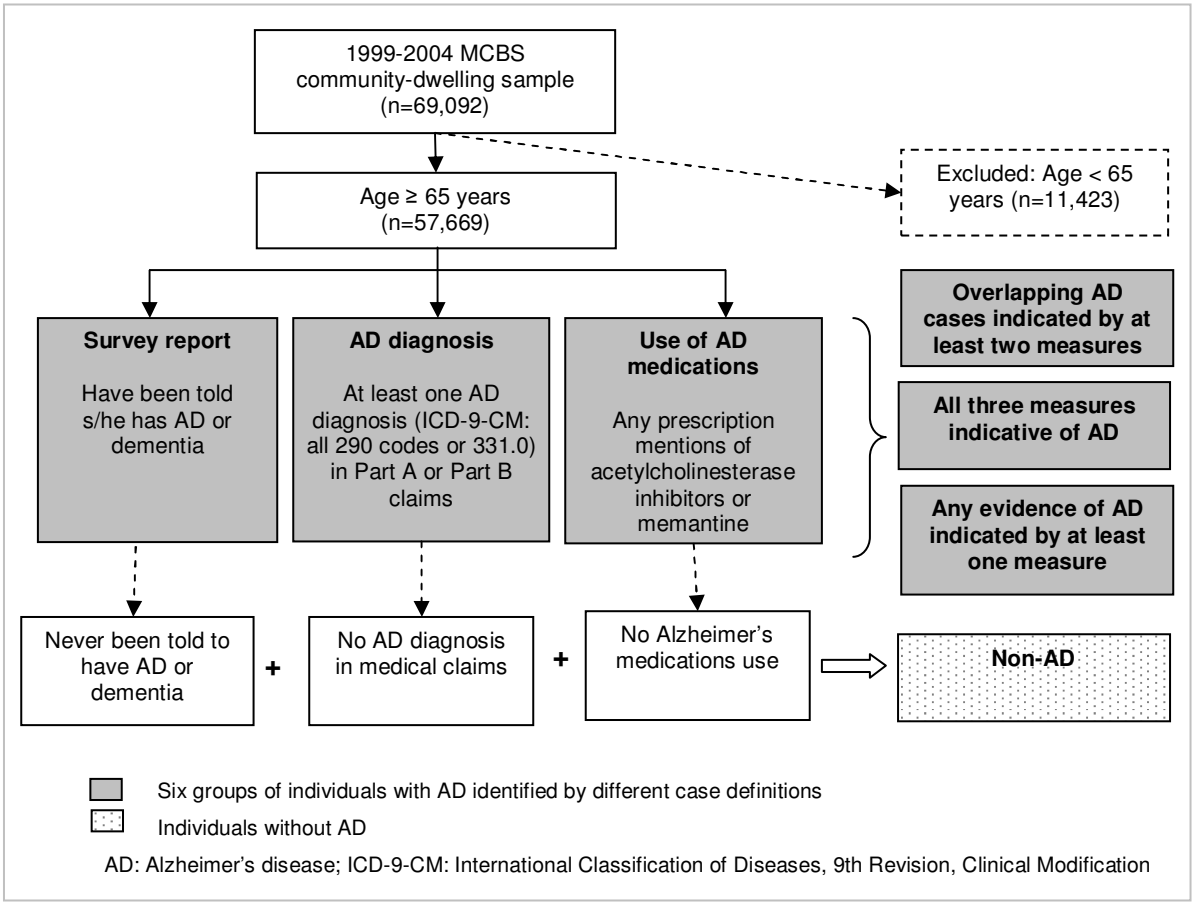


Figure 4.1: Sample extraction flowchart

Measures

Memory and cognitive problems, a core feature of AD, were assessed using two available variables in the MCBS: memory loss interfering with daily activities (yes/no) and decision-making problems interfering with daily activities (yes/no). Functional limitations as a critical part of the progression of AD are studied in most longitudinal and treatment studies [Mohs et al., 2000]. They are measured in the MCBS using six basic activities of daily living (ADLs) (i.e., bathing, dressing, eating, transferring in and out of chairs, walking and toileting), and six instrumental ADLs (IADLs) (i.e., using the phone, doing light housework, doing heavy housework, making meals, shopping and managing money). For each activity, survey respondents were asked whether they had any difficulty performing the activity, received help with the activity, or were unable to perform the activity because of health problems. For each

ADL, they also were asked whether they needed supervision with the activity. A survey subject was coded as impaired for an ADL or IADL if she or he needed any assistance performing or could not perform the activity. Three approaches were used to characterize the degree of functional limitations: 1) the number of ADL impairments and the number of IADL impairments (two separate, continuous variables); 2) dummy variables indicating impairments of specific ADLs and IADLs; and 3) Katz Index of ADLs, a hierarchical classification of individuals into eight mutually exclusive groups by type and number of ADL impairments [Katz et al., 1963]. The Katz Index incorporates continence but excludes walking as an ADL. The Katz index ranges from no impairments (level A) to total dependency on all six ADLs (level G), plus a category of "all other" to capture individuals not classified in levels A through G. This hierarchical classification allows a comparison between individuals with several specific impairments on ADLs and those without any impairments [Hill et al., 2006]. Additionally, general health status (coded as excellent, very good, good, fair and poor) also was studied.

Expenditure data in the MCBS were developed through a reconciliation process combining survey information and Medicare administrative files [CMS, 2003]. Personal total health care expenditures are defined as the MCBS aggregated payments across all types of services, not including rebates that may be paid by the manufacturer to the insurer. Prescription drug expenditures are imputed total payments in the MCBS for prescription medicines received from all sources, including Medicare, Medicaid, Medicare MCOs, private MCOs, the Veterans Administration, employer-sponsored private insurance, individually-purchased private insurance, private insurance from an unknown source, out-of-pocket payments and public health plans other than Medicare or Medicaid [CMS, 2003]. Detailed procedures and criteria used in the MCBS to impute missing payments for medical services and prescription drugs are discussed elsewhere [England et al., 1994]. Briefly, a full set of internally consistent expenditure and payment records were created from

these sources with minimal partial data discarded. The Consumer Price Index (CPI) for medical care [Bureau of Labor Statistics, 2007], which reflects the average price paid for a typical market basket of health care services, was used to adjust for inflation and to standardize all expenditures to 2007 dollars.

Analysis

We quantified the agreement between case definitions using raw agreement (%) illustrated by a Venn diagram and chance-corrected agreement measured by the kappa coefficient [Cohen, 1960]. Kappa coefficients were interpreted as follows: values of less than 0: poor agreement; 0-0.20: slight; 0.21-0.40:fair; 0.41-0.60: moderate; 0.61-0.80 substantial; and 0.80-1.00: almost perfect agreement [Landis and Koch, 1977]. McNemar's test was used to test the significance of discordance between two case definitions [McNemar, 1947].

The sensitivity (i.e., the rate of true positives) is the ability of the case definition to correctly identify AD cases, whereas the specificity (i.e., the rate of true negatives) is the ability to correctly identify those who do not have AD. Positive predictive value (PPV) represents the proportion of individuals with AD among the identified cases, whereas negative predictive value (NPV) represents the proportion of individuals without AD among those identified as not having AD [Szklo and Nieto, 2004]. These statistics provide a measure of the validity of the three case definitions, using one of them at a time as the "gold standard" definition for AD. Student's t-tests and chi-square tests were performed to assess the statistical significance of differences in characteristics among individuals in each of the six AD groups versus non-AD individuals. All analyses were conducted using SAS version 8.2 (SAS Institute, Cary, NC).

4.4 Results

Agreement between case definitions

A total of 3,198 individuals (5.55%) were classified as having AD using any of the three case definitions (Figure 4.2). Using survey report alone yielded more cases (n=1,994 or 3.46%) than did either diagnosis codes in medical claims (n=1,589 or 2.76%) or AD prescription drugs alone (n=1,160 or 2.01%). Of those with survey-reported AD, only 35.3% had an AD diagnosis in the medical claims in a given year and 32.8% took acetylcholinesterase inhibitors or memantine. Eleven hundred seventy-nine (2.04%) of the 57,669 elderly study sample were categorized as having AD according to at least two measures: 703 (1.22%) by survey report and diagnosis codes, 655 (1.14%) by survey report and pharmacologic treatments for AD, and 553 (0.96%) by diagnosis and pharmacologic treatments. Only 366 (0.63%) cases were indicated by all three definitions.

Chance-corrected agreement measured by kappa coefficients was low: survey-reported AD vs. AD diagnosis in claims records (kappa=0.37, 95% confidence interval [CI]: 0.35-0.39); survey report vs. use of Alzheimer's prescription medicine (kappa=0.40, 95% CI: 0.38-0.42); and AD diagnosis vs. use of Alzheimer's medications (kappa=0.38, 95% CI: 0.36-0.41). According to McNemar's tests, proportions of individuals with AD were significantly different between case definitions: survey report vs. diagnosis (χ^2=75.34, p<0.001); survey report vs. medication use (χ^2=377.20, p<0.001); and diagnosis vs. medication use (χ^2=112.02, p<0.001).

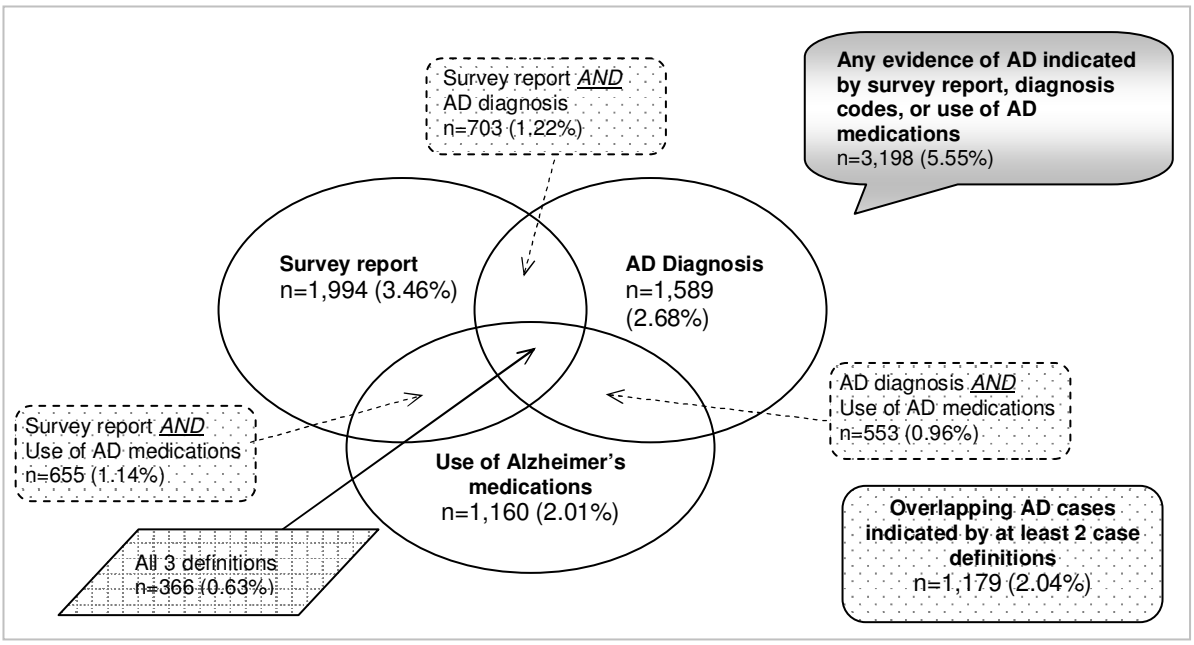


Figure 4.2: Venn diagram of raw agreement between classifications according to different definitions for identifying individuals with Alzheimer's disease (AD)

Sensitivity, specificity, positive and negative predictive values of case definitions

Using diagnosis in medical claims as a definitive measure for AD (i.e., "gold standard"), survey report correctly identified only 44.2% (i.e., the sensitivity) of individuals with AD, indicating that more than 55% of individuals with a diagnosis AD did not report this illness. The percentage of false negatives was even higher: 65%, among users of AD medications (sensitivity=34.8%). The PPVs were poor—only 35.3% truly had AD among the survey-reported cases and 47.7% among AD medications users. However, survey report and medication use did not falsely classify non-AD individuals as having AD (specificity=97.7% and 98.9%, NPV= 98.4% and 98.2%, respectively). If survey report had been used as the "gold standard", AD diagnosis and use of AD medicine were specific (98.4% and 99.1%) but had low sensitivity (32.3% and 32.8%). When changing the "gold standard" to medication use, survey report and diagnosis were more sensitive (56.5% and 47.7%) but slightly less specific (97.6% and 98.2%).

Table 4.1: Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of different case definitions for identifying individuals with Alzheimer's disease (AD)

Case definition	Number of AD cases under each gold standard	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Gold standard: AD diagnosis in medical claims	1589				
Survey report	703	44.2	97.7	35.3	98.4
Use of AD medications	553	34.8	98.9	47.7	98.2
Gold standard: survey report	1994				
AD diagnosis in medical claims	703	32.3	98.4	44.2	97.7
Use of AD medications	655	32.8	99.1	56.5	97.6
Gold standard: Use of AD medications	1160				
Survey report	655	56.5	97.6	32.8	99.1
AD diagnosis in medical claims	553	47.7	98.2	34.8	98.9

Prevalence of AD by patient characteristics

The prevalence of AD increases with age for each definition (Table 4.2). In the youngest age group (65 to 69 years), the proportions of AD ranged from 0.11% to 1.38%, compared with 1.21%-10.59% in the oldest age group (80+ years). As expected, using at least one case definition to identify individuals with AD produced estimates on the high end of the range; classification based on all three case definitions yielded estimates on the low end of the range. Using only a single definition, survey report produced slightly higher prevalence rates than did AD diagnosis, which in turn produced higher prevalence rates than did use of AD medications.

Table 4.2: Percentage with Alzheimer's disease (AD) by age group and by case definition

Age groups (yr)	Survey report (n=1,994)	AD diagnosis (n=1,589)	Use of AD medications (n=1,160)	Any evidence of AD (n=3,198)	At least two definitions (n=1,179)	All three definitions (n=366)
65-69 (n=12,637)	0.91	0.56	0.33	1.38	0.32	0.11
70-74 (n=13,606)	1.71	1.17	1.06	2.67	0.92	0.35
75-79 (n=11,697)	2.88	2.35	2.00	4.89	1.78	0.56
80+ (n=19,729)	6.64	5.49	3.75	10.59	4.09	1.21
Total (n=57, 669)	3.46	2.76	2.01	5.55	2.04	0.63

Individuals with AD were on average 82 years old, and were significantly older than their peers without AD ($p<0.001$) (Table 4.3). Individuals with AD were predominately female (>60%) and white (>80%). Individuals meeting all three case definitions and those with survey-reported AD had the highest rates of proxy use—64.5% and 63.5%, respectively. Individuals identified by use of Alzheimer's medications had the lowest rate of proxy use (43.0%) among the six case definition groups. Proxy use was even lower, only 11.2%, in the non-AD population. The impact of AD also is reflected by poorer general health, memory and decision-making skills relative to individuals without such illness. Among individuals with AD, 37.9%-47.5% rated themselves in fair or poor health, compared with 22.0% in the non-AD population ($p<0.001$). On average, 57.4%-84.9% and 50.1%-75.7% of individuals with AD had memory loss and decision-making problems, respectively, to the point that they interfered with daily activities, whereas only 8.1% and 4.5%, respectively, of the non-AD beneficiaries reported these symptoms ($p<0.001$).

Table 4.3: Patient characteristics by case definition identifying individuals with Alzheimer's disease (AD)

Characteristics	Survey report (n=1,994)	AD diagnosis (n=1,589)	Use of AD medications (n=1,160)	Any evidence of AD (n=3,198)	At least two definitions (n=1,179)	All three definitions (n=366)	Non-AD (n=54,471)
Mean age (s.d.), year	82.1 (7.5)	82.3 (6.9)	81.3 (6.3)	81.9 (7.2)	82.2 (6.7)	81.6 (6.6)	76.0 (7.3)
Male, %	38.3	36.5	38.8	37.7	37.2	40.2	43.2
Race/ethnicity, %							
White, not Hispanic	81.1	82.4	86.3	82.7	83.0	83.9	86.4
Black, not Hispanic	13.0	12.6	9.1	11.8	12.4	11.5	8.8
Hispanic	2.2	2.5	2.8	2.2	2.5	3.8	2.4
Other	3.7	2.5	1.8	3.3	2.1	0.8	2.4
Proxy respondent, %	63.5	53.2	43.0	51.6	61.3	64.5	11.2
General health status, %							
Excellent	8.0	8.3	8.9	8.7	7.8	6.6	15.9
Very good	16.6	17.4	18.5	17.6	16.7	16.5	28.8
Good	27.9	29.0	34.7	29.7	30.8	29.4	33.3
Fair	28.3	26.9	26.7	27.0	27.0	32.7	16.6
Poor	19.2	18.4	11.2	17.0	17.7	14.8	5.4
Memory loss interfering with daily activities, %	74.9	57.4	64.2	60.1	77.9	84.9	8.1
Decision-making problems interfering with daily activities, %	67.8	50.1	52.6	52.0	69.4	75.7	4.5

s.d.: standard deviation

Functional status

Functional limitations were much more common among individuals with AD, regardless of the case definition employed, than in individuals identified as not having this condition (Table 4.4). Individuals meeting any of the three case definitions had an average of 1.4 ADL impairments, compared with 0.2 ADL limitations in those without AD. Across all AD groups, bathing and dressing were the two most prevalent ADL impairments, and eating was the least prevalent. In general, ADL impairments were more pronounced in individuals identified by survey report and less prevalent in users of AD medications. The same trend was reflected by the Katz hierarchical index: 35.7% of individuals with survey-reported AD were fully independent in all ADLs (level A), compared with 51.1% of individuals identified by

use of AD medications. More than 15% of individuals with survey-reported AD were impaired in five or all six ADLs (levels F or G), compared with 8.4% of AD medicine users.

Performing IADLs is more cognitively demanding than is ADLs. Hence, IADL limitations were even more pronounced in individuals with AD. On average, individuals meeting all three case definitions had 2.6 IADL impairments, compared with 0.6 IADL impairments reported by individuals without AD ($p<0.001$). Of the six IADL items assessed, individuals with AD were more likely to receive help with or fail to perform heavy housework due to health problems, but were less likely to have difficulty using the telephone. Nearly half of the individuals with any of the three case definitions reported difficulty or inability to make meals (42.8%), shop (49.9%) and manage money (49.8%), each of which requires higher levels of cognitive functioning than doing housework. Similar to the trend observed for ADLs, individuals with survey-reported AD were more likely to have IADL impairments relative to other AD groups, whereas those identified by medications were less likely to report IADL limitations.

Table 4.4: Functional health status by case definition for Alzheimer's disease (AD)

Functional status	Survey report AD diagnosis (n=1,994)	(n=1,589)	Use of AD medications (n=1,160)	Any evidence of AD (n=3,198)	At least two definitions (n=1,179)	All three definitions (n=366)	Non-AD (n=54,471)
Number of ADL impairments, mean (s.d.)	1.7 (2.1)	1.4 (2.0)	1.0 (1.7)	1.4 (2.0)	1.6 (2.1)	1.6 (2.0)	0.2 (0.7)
ADL limitation, %							
Bathing	44.5	37.1	27.9	35.3	56.8	44.3	5.3
Dressing	36.0	30.1	22.0	28.3	37.8	37.2	3.7
Eating	13.6	12.4	6.5	10.7	13.4	11.8	0.7
Transferring in and out of chairs	27.9	23.4	16.2	22.4	25.7	26.0	2.9
Walking	27.5	21.7	15.6	21.9	24.3	24.0	3.5
Toileting	21.6	18.7	12.2	17.0	21.3	20.2	1.4
Katz index of ADLs, %							
A: independent in all activities	35.7	42.3	51.1	43.5	38.3	36.9	73.1
B: independent in all but one activity	21.0	21.3	22.3	21.8	20.3	21.0	22.0
C: dependent in bathing and one other activity	7.3	6.6	5.6	6.4	7.2	7.1	1.5
D: dependent in bathing, dressing and one other activity	6.6	5.0	4.5	5.0	6.5	6.8	0.8
E: dependent in bathing, dressing, going to toilet and one other activity	2.6	2.6	1.7	2.3	2.6	2.7	0.3
F: dependent in bathing, dressing, going to the toilet, transferring and one other activity	8.0	6.0	4.9	6.2	7.1	8.8	0.6
G: dependent in all activities	7.7	7.1	3.5	6.0	7.8	6.0	0.3
Other: dependent in at least two activities and not classified in A-G	11.1	9.1	6.4	8.8	10.2	10.7	1.4
Number of IADL impairments, mean (s.d.)	3.2 (2.3)	2.7 (2.4)	2.4 (2.3)	2.6 (2.4)	3.2 (2.3)	3.4 (2.3)	0.6 (1.3)
IADL limitation, %							
Using the phone	33.0	26.0	21.4	25.1	32.8	36.3	3.1
Doing light housework	48.6	40.8	33.1	40.2	45.5	48.9	9.5
Doing heavy housework	63.5	58.5	52.6	56.8	63.2	66.4	24.4
Making meals	53.2	43.4	37.3	42.8	51.3	57.1	7.2
Shopping	60.5	50.4	44.6	49.9	59.2	63.9	12.2
Managing money	63.2	49.9	47.0	49.8	63.7	69.1	5.6

ADL: activity of daily living; IADL: instrumental activity of daily living; s.d.: standard deviation

Expenditure variations by case definition

We plotted average total health care expenditures (Figure 4.3) and prescription drug expenditures (Figure 4.4) against the various case definitions. Regardless of the definition employed, individuals with Alzheimer's disease, on average, had significantly higher total and prescription drug expenditures compared with Medicare beneficiaries without AD. The mean total expenditures were 1.6- to 2.4-times higher in those with AD than in the non-AD population ($16,547-$24,937 vs. $10,371, respectively, p<0.001). Individuals with AD diagnoses in their medical claims were very costly and produced mean total expenditure estimates on the high end of the range, whereas those with survey-reported AD were on the low end. Drug expenditures for individuals with AD were on average 1.4- to 2.1-times higher than beneficiaries without AD ($2,303-$3,502 vs. $1,681, respectively, p<0.001). Drug expenditures were approximately 40% higher in users of Alzheimer's medications ($3,298) compared with individuals identified by survey report ($2,303) and by AD diagnoses ($2,357).

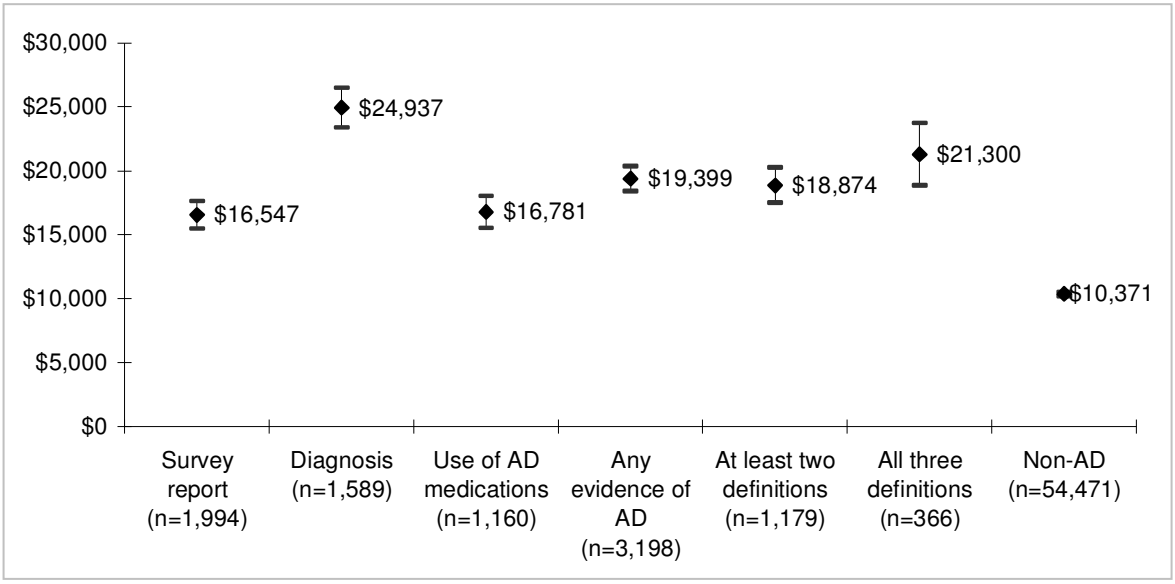


Figure 4.3: Total health care expenditures by Alzheimer's disease (AD) case definition (mean and 95% confidence interval, 2007$)

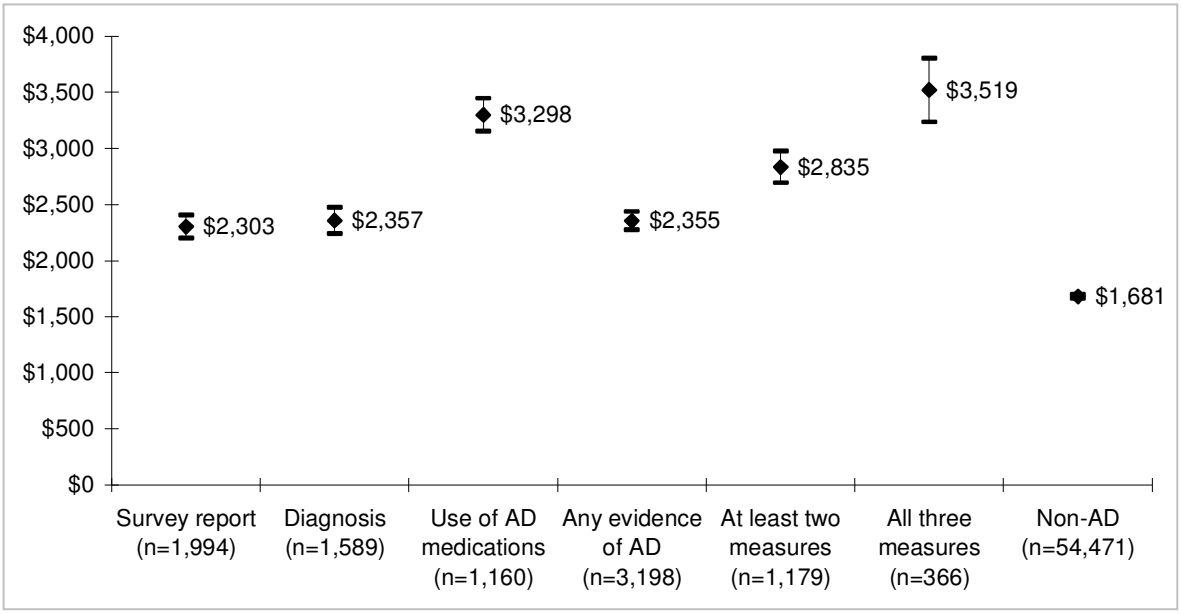


Figure 4.4: Prescription drug expenditures by case definition of Alzheimer's disease (mean and 95% confidence interval, 2007$)

4.5 Discussion

Our investigation into identifying individuals with AD in the MCBS showed that prevalence of and expenditure estimates associated with AD varied widely by the definition used for case determination. Use of any single case definition in observational studies may fail to capture the actual burden of AD at the population level. Therefore, one should exercise caution in interpreting current cost-of-illness studies and in applying these estimates to policy initiatives. Because different case definitions may select different subsets of individuals with AD, incorporating all sources of cases definitions available to researchers for case ascertainment is a crucial first step to assessing health care needs for individuals with AD.

Using a single case definition, such as an AD-associated diagnosis code in administrative data alone, to define individuals with AD has been shown to introduce errors of omission and commission [Newcomer et al., 1999; Fillit, 2000; Rice et al., 2001]. In our sample, survey report appeared to be the best single-definition approach, given the smallest number of missed cases: 505 cases with an AD diagnosis in medical claims or 886 users of

Alzheimer's medications. Relying on a diagnosis in claims data alone would have missed 1,291 survey-reported cases or 607 medication-identified cases who had no AD diagnosis in their medical claims. Using pharmacy data alone would even more seriously under-estimate the burden of AD—1,339 survey-reported cases or 1,036 diagnosed cases would have been missed. If an AD diagnosis in medical claims were used as a "gold standard," survey report and use of Alzheimer's medications were highly specific (i.e., low false-positive rates) but not sensitive (i.e., high false-negative rates). When we changed the "gold standard" to either survey report or medication use, our case definitions all yielded high specificities, which is not surprising in that the prevalence of AD is relatively low in the population. However, low sensitivities suggest that use of a single case definition may fail to capture all individuals with AD.

An analysis of the Medicare Alzheimer's Disease Demonstration, a longitudinal sample of individuals known to have some form of dementia, showed that claim-based AD prevalence rates were substantially underestimated because of underreporting of chronic conditions [Newcomer et al., 1999]. Fewer than 20% of the sample were classified as having dementia of the Alzheimer type (ICD-9-CM code: 331.0) based on a single year of claims, whereas 68% actually had an AD diagnosis from a referring physician at the time of entering the demonstration. Moreover, observers complain that many mild AD cases are underdiagnosed or uncoded in claims data because of social resistance to ageism and disabling stigma [Fillit, 2000; Rice et al., 2001]. The choice of diagnosis code also may be influenced by reimbursement scheme, which provides little financial incentive for coding AD as the primary diagnosis and instead encourages coding comorbidities, such as aspiration pneumonia as the primary diagnosis, to enhance reimbursement [Newcomer et al., 1999; Fillit, 2000; Brummel-Smith, 2001; Rice et al., 2001]. In our data, only 35% of individuals with survey-reported AD and 48% among users of AD medications had an Alzheimer's diagnosis listed as any diagnosis in their Medicare claims. Our findings support the evidence that

claims data only identify a subset of individuals with AD and that different case definitions produce widely divergent prevalence estimate [Fillit, 2000; Rice et al., 2001; Bloom et al., 2003; Pressley et al., 2003]. The disparity is, in part, attributable to the lack of a uniformly accepted definition or a gold-standard diagnostic test to identify Individuals with AD, which, in turn, poses a challenge to assess accurately the actual burden of AD. In order to enhance case ascertainment, researchers have suggested supplementing diagnoses in claims data with procedure codes, survey report, medical record reviews and direct physician reports of AD [Newcomer et al., 1999]. Our findings reinforce the use of survey report and suggest that pharmacy data also may be a valid supplement to claims records for case determination.

Using survey report, Kane and Atherly estimated that the prevalence per 100 population for Alzheimer's or other dementia was 1.30 for individuals 65-74, 5.75 for those 75-84, and 19.88 for individuals 85 and older; the overall prevalence rate was 6.8 for women and 3.93 for men in the 1991-1995 MCBS [Kane and Atherly, 2000]. Hill et al. [2006] estimated that, in the 1995-1998 MCBS, individuals with survey-reported AD (n=3,138, including community and facility residents) were equivalent to between 1.8 and 2 million Medicare beneficiaries in each of the four calendar years in the study period, which is consistent with estimates in epidemiological studies [Brookmeyer et al., 1998; US GAO, 1998]. Our analysis showed that, using only survey report, 1,994 (3.46 cases per 100 population) were classified as having AD: 1.32 for the 65-74 age group, 3.88 for the 75-84 age group and 8.69 for those age 85 and older, equivalent to between 0.77 to 1.18 million Individuals with AD annually during 1999-2004. Our estimates compare closely to those reported in the literature.

Using Medicare claims records, Weiner and colleagues estimated that per capita Medicare expenditures for individuals with dementia of the Alzheimer type (ICD-9-CM code 331.0) in 1992 fiscal year were $9,300 (2007$), almost twice as much as the expenditures for all beneficiaries [Weiner et al., 1998]. Hill et al. used Medicare MCO claims data for a

two-year period and reported 1.6-times higher annual costs for individuals with AD and related dementias (ICD-9-CM codes: all 290 codes, 797, 292.82, 291.2, 294.1, 294.8, and 331.0-331.2 or/and donepezil users) than those without such illness ($15,015 vs. $9,226, 2007$), after controlling for age, sex and comorbidities [Hill et al., 2002]. Taylor and Sloan [2000] argued that using only one or two years of claims data to identify AD cases may underestimate prevalence and bias cost estimates. Using the 1994 National Long Term Care Survey merged with 12 years (1984-1995) of claims, 3.1% of the Medicare population (including both community and institutional population) were classified as having AD (ICD-9-CM code: 331.0). On average, individuals with AD had 2.6-times higher annual health care costs compared to individuals without AD ($9,588 vs. $3,678, age- and sex-adjusted 2007$). Our study showed that using AD-associated diagnosis codes in a single year of claims data produced the highest mean total health care costs ($24,905), or 2.4-times the mean costs of non-AD individuals in the community. However, the total health care expenditures varied by as much as 50%, depending on the definition used for case identification. Therefore, one should be aware of methodological differences in calculating costs in the literature, especially when applying these estimates to AD policies [Bloom et al., 2003].

In interpreting the results presented here, several notes of caution are important. One involves the completeness of administrative claims records in the MCBS. Because Medicare MCOs do not submit claims with diagnoses to the CMS, Medicare does not have a record of covered or noncovered services provided to beneficiaries in these plans [CMS, 2003]. As a result, it is possible that conditions occurring during the group health coverage period may not be reflected in Medicare claims provided in the Cost and Use files. This may lead to underestimating the proportion of individuals with AD covered by Medicare using ICD-9-CM diagnosis codes in the claims data. Therefore, we performed subgroup analyses excluding managed care enrollees to verify the robustness of our findings (results not shown). In

general, excluded individuals with group health participation (n=10,204) were found somewhat younger and healthier, and had lower average health care costs. Twenty-seven hundred seventy-nine (5.85% vs. 5.55% for all) of those not enrolled in managed care were classified as having AD using any of three case definitions. The results throughout the subgroup analyses were consistent with our findings for the entire sample composed of a mix of fee-for-service and MCO enrollees. Total health care costs and drug expenditures varied by as much as 41% and 49%, respectively, depending on the definition used for case determination. Therefore, our findings are not sensitive to MCO enrollment status.

Another issue concerns the quality of self versus proxy report. Our data showed that 11% of non-AD beneficiaries' interviews were conducted by proxy, whereas 43%-64% (depending on the case definitions) of proxy response occurred on behalf of beneficiaries with AD. Among those with survey-reported AD, proxy-reported cases had poorer health and more functional impairments than self-reported cases as one would expect in that proxy was used if the sample person was unable to respond due to physical or mental problems. Although self-report may have greater errors in individuals with cognitive impairment relative to the general population, it is not clear whether proxy reports are systematically more or less accurate than self reports. Using an AD diagnosis in medical claims as a "gold standard," we recalculated the sensitivity and specificity of survey-reported AD by proxy use. Self report appeared more specific (98.8%) then proxy (88.7%), but much less sensitive (24.8% vs. 61.4%), indicating that individuals with a diagnosis of AD may be more likely to under-report this condition than proxy respondents. Some studies have shown that individuals with dementia may overestimate their cognitive and functional abilities compared with proxy ratings [Koss et al., 1993; Tierney et al., 1996; Farias et al., 2005]. On the other hand, a concern has been raised that greater caregiver burden may lead to overstated functional impairments of individuals with cognitive impairment [Rothman et al., 1991; Long et al., 1998]. A review of the validity of proxy respondents for older persons with cognitive impairment and

other chronic debilitating conditions found that proxies tend to describe more functional impairment and poorer emotional well-being than do patients themselves [Neumann et al., 2000]. Additionally, proxy ratings of ADL and IADL disability were found to increase with caregiving hours and higher perceived caregiver burden. Our data showed that proxy use was higher in individuals with survey-reported AD, who also had higher prevalence of memory loss and decision-making problems and poorer functional status. However, we were not able to determine whether proxy respondents rated these impairments more negatively than they truly were. In the future it will be important to investigate the validity of proxy reports in order to accurately assess the cognitive ability and functional status in individuals with dementia.

A third limitation pertains to the accuracy of prescribed medicine events in the MCBS. Information on prescription drug use was ascertained from survey data rather than from actual pharmacy claims (which were not paid by Medicare). Survey respondents were asked to bring to the interview bottles, tubes and prescription bags provided by the pharmacy as a way to verify survey-reported medication use. However, the prevalence of AD still may be underestimated due to under-reporting of medication use by a demented patient or by a proxy respondent. Prescriptions filled several months earlier may not be a salient event in the typical respondent's memory [England et al., 1994]. Individuals may have a difficult time saving all prescription containers (especially prescription bags) over the typical four-month span between interviews. Furthermore, it is noteworthy that an estimated 50% of individuals with AD are diagnosed, and only half of those diagnosed are being treated [Evans, 1990]. Therefore, pharmacy data should be used only as a supplement definition for case ascertainment rather than a definitive measure for AD.

Despite these concerns, we believe that our analyses convey important information on methodological improvements for estimating the prevalence and costs of AD, and hope that they provide greater confidence in case ascertainment using secondary data. Future

research should seek to evaluate the performance of other supplements to medical claims data, such as cognitive function assessment, apolipoprotein E genotype testing (i.e., a genetic test for Alzheimer risk gene) and medical record review and longitudinal data, for case ascertainment.

4.6 Conclusions

Alzheimer's disease is set to become one of the developed world's largest health care burdens over the coming decades. Current estimates using only a single case definition may fail to capture the full burden of AD at the population level. Thus, one should exercise caution in interpreting current cost-of-illness studies and in applying these estimates to policy initiatives. Future studies using secondary data to assess the health care needs for individuals with AD should first develop a comprehensive definition to identify a broad spectrum of cases, proceeding to exclusion of false-positives if warranted.

CHAPTER 5

THE CONCENTRATION AND PERSISTENCE OF HEALTH CARE EXPENDITURES AND PRESCRIPTION DRUG EXPENDITURES IN MEDICARE BENEFICIARIES WITH ALZHEIMER'S DISEASE

5.1 Abstract

Introduction: Health care expenditures in Medicare are highly concentrated in a small proportion of beneficiaries. The purpose of this study was to quantify the concentration and persistence of overall and prescription drug expenditures in individuals with Alzheimer's disease (AD) and to determine the person-level characteristics associated with future expenditure levels.

Methods: Data were obtained from the 1999-2004 Medicare Current Beneficiary Survey linked with Medicare claims. Elderly, community-dwelling individuals with AD were rank-ordered by overall and drug expenditures. The proportion of expenditures accounted for by the top 10%, top 25% and top 50% of individuals was calculated. A transition probability matrix was used to illustrate the change in expenditure percentiles from one year to the next. Ordered logit models incorporating prior expenditure, the Charlson Comorbidity Index, functional status and other background covariates were estimated to predict the level of subsequent-year expenditures.

Results: Average per capita health care expenditures in our AD sample ranged from $19,927 to $22,144, and prescription drug expenditures ranged from $1,797 to $2,943 during the study period. Individuals in the top 10% of the expenditure distribution accounted for 38%-47% of overall health expenditures and incurred 31%-36% of overall drug expenditures. A quarter of the highest-spending 10% for total health expenditures remained in the top

decile in the next year, whereas 21% of them moved to the bottom half of the distribution in the subsequent year. Half of the highest 10% with drug expenditures retained this ranking but only 9% became the bottom 50% in a second year. Prior expenditures and Charlson comorbidity scores, but not functional status, were strong predictors of the level of future expenditures.

Conclusions: Overall health care and drug expenditures were highly concentrated and persistent over a two-year period in this AD population. Prescription drug expenditures exhibited less concentration but more persistence than did overall health expenditures. Results from this study may further our knowledge of how expected high expenditures in individuals with AD may be reduced with improved care coordination and effective disease management.

Keywords: Alzheimer's disease, high health expenditures, drug expenditures, Charlson Comorbidity Index, Medicare

5.2 Introduction

Alzheimer's disease (AD) is a neurodegenerative disease characterized by a progressive deterioration in cognition, memory and functional ability. Individuals with AD often experience prominent and multiple symptoms, including psychological symptoms and behavioral disturbances, that are distressing and place substantial burdens both on the individuals themselves and the health care system [Grossberg, 2002; Fillit and Hill, 2004; Kaufer et al., 2005]. According to a recent report by the Alzheimer's Association [2007], 13% of individuals age 65 and older are living with AD; nearly half of persons older than 85 are affected by this disease. A review of 21 cost-of-illness studies using 1985-2000 data found that inflation-adjusted costs ranged from $5.6 to $88.3 billion nationally [Bloom et al., 2003]. In 2000, individuals with AD represented less than 5% of Medicare beneficiaries [Taylor and

Sloan, 2000], while they accounted for 14.4% of overall Medicare spending [Alzheimer's Association, 2001]. Individuals with both AD and comorbid conditions have been shown to accrue substantially higher expenditures compared with the matched control subjects without AD [Hill et al., 2002]. However, little is known about the extent to which high and persistent health expenditures are systematically associated with individual characteristics and prior utilization. Despite numerous cost-of-illness studies on AD, there is insufficient understanding about the characteristics of individuals with AD who have high expenditures, and who continue to spend a disproportionately large share of money. Identifying these high-expenditure and persistent users is crucial for managed care organizations (MCOs) seeking to target cases for intensive medical and financial management and to improve care coordination [Russell and Chaudhuri, 1992; Ash et al., 2001].

Health care expenditures in the U.S. are highly concentrated in a small proportion of population [Conwell and Cohen, 2005]. In a nationally representative sample from the 1996 Medical Expenditure Panel Survey (MEPS), the highest-spending 1% of individuals made up 27% of overall health expenditures and the top 5% accounted for 55% of overall expenditures, whereas the bottom 50% accounted for only 3% of overall expenditures [Berk and Monheit, 2001]. Based on a longitudinal population of Medicare enrollees during 1987-1995, the top 1% of individuals in the expenditure distribution accounted for nearly 20% of overall expenditures, and the proportion of expenditures attributable to this group increased over time [Garber et al., 1997]. In a managed care setting, Meenan et al. [2003] demonstrated that the most expensive 0.5% and 1% of cases represented 20% and 29%, respectively, of total expenditures.

This consistent trend has motivated researchers to investigate the extent to which individuals in high expenditure percentiles persist from one year to the next. Evidence has shown that Medicare beneficiaries who incur high expenditures in one year tend to have higher-than-average expenditures in other years [Garber et al., 1997]. Based on the

1996-2000 MEPS data, 31% of non-institutionalized Medicare beneficiaries whose total expenditures were in the top 10% in year 1 retained this ranking in year 2; 49% of those beneficiaries ranked in the top 25% remained in the top 25%, and 72% of those in the top 50% stayed in the top 50% in the next year [Banthin and Miller, 2006]. Previous studies have evaluated the persistence of overall health expenditures using administrative claims or survey data for the elderly with public insurance [Garber et al., 1997; Kapur et al., 2000], individuals with private insurance [Russell and Chaudhuri, 1992; Meenan et al., 1999; Meenan et al., 2003], a nationally representative sample of children [Liptak et al., 2006] and a national sample of the civilian, noninstitutionalized population [Monheit, 2003; Cohen et al., 2006]. Results from these studies generally support the notion that high-expenditure individuals account for a disproportionate share of health care expenditures and that a small group exhibits persistently high expenditures.

As with total health expenditures, drug expenditures also exhibit persistence over time, especially in the elderly population. Two studies conducted during the early 1990s examining the persistence in drug expenditures in the Pennsylvania Pharmaceutical Assistance Contract for the Elderly program [Stuart et al., 1991; Coulson and Stuart, 1992] suggested a strong degree of persistence in prescription drug expenditures by the elderly, particularly among the heaviest users. In Medicare, drug expenditures are highly persistent. Wrobel et al. [2003] found that demographic variables in the 1999 Medicare Current Beneficiary Survey explained only 5% of the variation in drug expenditures in 2000, whereas adding health status measures increased the explanatory power (i.e., adjusted R^2) to 10%-24%. Incorporating the prior year's drug expenditures raised explanatory power to 55%. Banthin and Miller [2006] showed that Medicare prescription drug expenditures were even more persistent than were total expenditures. Among Medicare beneficiaries whose drug expenditures were in the top 10% in year 1, 55% remained in the top decile in year 2; 67% of those ranked in the top 25% retained this ranking, and 81% of those in the top 50% retained

this ranking in the next year. The predictive model including information on drug treatment categories and prior drug expenditures explained nearly 53% of the variation in subsequent-year drug expenditures.

According to the Department of Health & Human Services Administration on Aging [2006], the number of individuals age 65 and older is projected to increase from 40.2 million in 2010 to more than 71.5 million by 2030. It is likely that the health care needs for individuals with AD will continue to rise, as advancing age is the single most important risk factor for AD. Unlike studies of the general Medicare beneficiaries, we sought to quantify the concentration and persistence of overall and prescription drug expenditures in individuals with AD, a population with progressive disabilities and high expenditures. Specifically, the present study was designed to examine whether high-expenditure users account for a disproportionate share of expenditures among individuals with AD, to determine the extent to which the individuals in high-expenditure percentiles persist from one year to the next, and what characteristics are associated with expenditure persistence.

5.3 Methods

Data source

Data for this study were obtained from the 1999-2004 waves of the Medicare Current Beneficiary Survey (MCBS) Cost and Use files, which are nationally representative and comprehensive data sets of health status survey, health care use and expenditure data for the Medicare population. The MCBS is drawn from Centers for Medicare and Medicaid Services' (CMS) Medicare enrollment file. The survey data include information on the cost and use of all types of medical services, and can be linked directly with Medicare claims. Available Medicare Part A (i.e., hospitalization, skilled nursing facility, hospice, and home health care) and Part B (i.e., physician visits, specified outpatient care and some outpatient medications) claims records include diagnosis codes, utilization, charges and reimbursement

for all medical services [CMS, 2003]. The survey oversamples significant subpopulations (e.g., older beneficiaries) allowing increased statistical power for subgroup analyses, such as for elderly persons with AD. Each individual, or an appropriate proxy respondent (usually a family member or close acquaintance), is interviewed three times per year during a four-year period to form a continuous profile of the individual's health care experience [CMS, 2006b]. The survey features a longitudinal rotating panel design, in which one-fourth of the panel is replaced by new subjects each year, resulting in a cumulative sample size of three 4,000-beneficiary cohorts or approximately 12,000 beneficiaries in any given year. Therefore, the data can support both cross-sectional (e.g., concentration of expenditures) and longitudinal analyses (e.g., persistence of expenditures).

Sample

The study sample was drawn from community-dwelling Medicare beneficiaries, defined as individuals who were not institutionalized for more than 90 days at a time during a year [CMS, 2006c]. Survey interviews were completed by either sample members or proxy respondents if the sample member was unable to respond due to physical or mental problems. The inclusion criteria restrict the sample to community-dwelling beneficiaries aged 65 years or older (Figure 5.1). To identify individuals with AD, we incorporated three case definitions. Use of a single case definition, such as AD diagnosis in administrative claims data alone, has been shown to introduce errors of omission and commission [Newcomer et al., 1999; Fillit, 2000; Rice et al., 2001; Pressley et al., 2003; Lin, 2008a]. Therefore, we defined individuals who met any of the following case definitions as having AD:

1) affirmative answer to the question "Has a doctor ever told you that you had Alzheimer's disease or dementia?"; or

2) at least one AD diagnosis, defined by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes: all 290 codes (senile and presenile organic psychotic conditions) and 331.0 (AD) in Medicare Part A or Part B claims files; or

3) use of any Alzheimer's medications, including acetylcholinesterase inhibitors (i.e., donepezil [Aricept®], rivastigmine [Exelon®], and galantamine [Reminyl® or Razadyne®]) and memantine (Namenda®). These medications were identified by the drug names in survey-reported data, i.e., pharmacy administration data were not used.

Individuals with AD who had any managed care participation during that year were excluded because these Medicare MCOs do not submit claims to the CMS. As a result, Medicare does not have a record of covered or noncovered services provided to beneficiaries in these plans [CMS, 2003]. Thus, a total of 2,779 observations from the 1999-2004 waves of the MCBS were pooled for the expenditure concentration analysis. Of these, there were 1,861 eligible, unique individuals. We excluded persons with only one year of claims (including deaths of sample respondents and nonresponse in later rounds). We then retained individuals who had medical claims data for both year *t* and year *t+1*, and excluded data for year *t+2* among those who were observed more than twice in the data set. As a result, two-year panels of 671 unique individuals with AD were analyzed for expenditure persistence. Of these, 155 (23%) were in the 1999-2000 cohort, 102 (15%) were in the 2000-2001 cohort, 136 (20%) were in the 2001-2002 cohort, 147 (22%) were in the 2002-2003 cohort, and 131 (20%) were in the 2003-2004 cohort; seventy-nine individuals (11.8%) died during the second year.

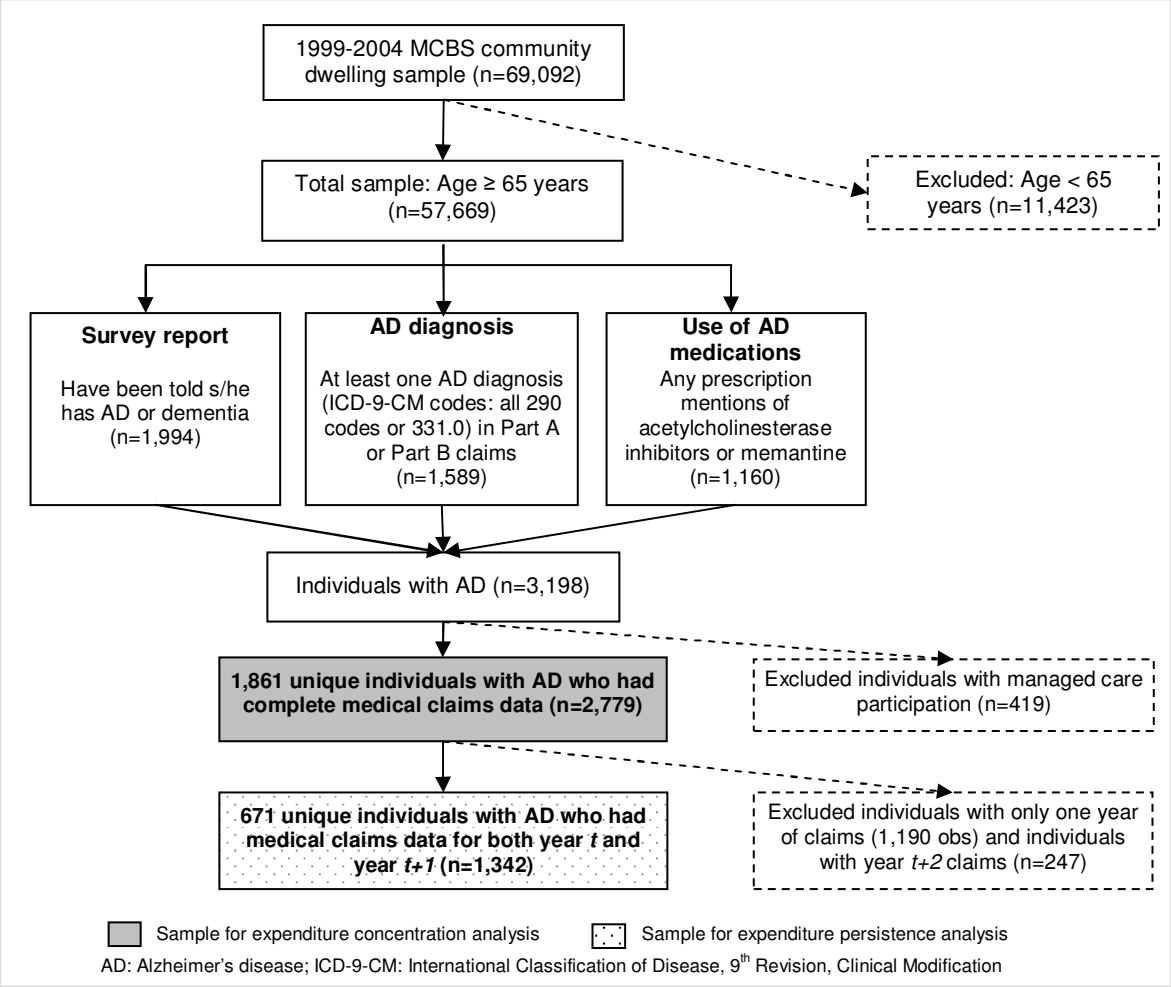


Figure 5.1: Sample extraction flowchart

Measures

Expenditure data in the MCBS were developed through a reconciliation process combining survey information and Medicare administrative files [CMS, 2003]. Personal health care expenditures are defined as the standardized payments aggregated across all types of services, not including rebates that may be paid by the manufacturer to the insurer. Drug expenditures are imputed total payments for any prescription drugs from Medicare, Medicaid, Medicare MCOs, private MCOs, the Veterans Administration, employer-sponsored private insurance, individually-purchased private insurance, private insurance from an unknown source, out-of-pocket payments and public health plans other than Medicare or Medicaid [CMS, 2003]. Detailed procedures and criteria used in the MCBS to impute missing

payments for medical services and prescription drugs are discussed elsewhere [England et al., 1994]. Briefly, a full set of internally consistent expenditure and payment records were created from these sources with very little partial data discarded. All expenditures were converted into constant 2007 dollars using the Consumer Price Index for medical care [Bureau of Labor Statistics, 2007], which reflects the average price paid for a typical market basket of health care services.

We examined the top five therapeutic classes for prescription medicines, which are designated by First Data Bank [CMS, 2007a], used by the top 10%, top 25%, top 50%, and bottom 50% of the individuals with AD; drug expenditures included expenditures for all drugs, however. We incorporated background covariates in addition to prior expenditure levels for regression analyses. Health status measures include perceived general health (categorized as excellent or very good, good and fair or poor), memory loss (yes/no) and decision-making problems interfering with daily activities (yes/no). Six basic activities of daily living (ADLs) (i.e., bathing, dressing, eating, transferring in and out of chairs, walking and toileting) and six instrumental ADLs (IADLs) (i.e., using the phone, doing light housework, doing heavy housework, making meals, shopping and managing money) were used to measure functional status. An individual was coded as impaired for an ADL or IADL if she or he needed any assistance performing or could not perform the activity. We tested three different specifications to characterize the degree of functional limitations: Model 1 used the number of ADL impairments and the number of IADL impairments (two separate, continuous variables); Model 2 used dummy variables indicating impairments of specific ADLs and IADLs; and Model 3 used Katz Index of ADLs, a hierarchical classification of individuals into eight mutually exclusive groups by type and number of ADL impairments [Katz et al., 1963]. The Katz Index incorporates continence but excludes walking as an ADL. The index ranges from no impairments (level A) to total dependency on all six ADLs (level G), plus a category of “all other“ to capture individuals not classified in levels A through G. This hierarchical

classification allows a comparison between individuals with several specific ADL impairments and those without any impairments [Hill et al., 2006].

We used the Charlson Comorbidity Index (CCI) with the Deyo modification [Charlson et al., 1987; Deyo et al., 1992] to control for the effects of comorbid conditions on overall and drug expenditures. The CCI for a given year was derived from the presence of various diagnoses (defined by the ICD-9-CM codes) in the medical claims, including claims for inpatient, outpatient, physician visit, skilled nursing facility, home health and hospice service. The CCI comprises 17 comorbidity categories; each condition is assigned a weight as 1, 2, 3, or 6, reflecting the magnitude of the adjusted relative risks associated with that comorbidity. The weights for all conditions are accumulated to calculate the CCI scores, with higher numbers representing a greater burden of comorbidity.

Analysis

The study sample was rank-ordered by overall expenditures and by drug expenditures, and then categorized into four groups based on their percentile position in the expenditure distribution of a given year: 1) top 10%, 2) next 11%-25%, 3) next 26%-50%, and 4) bottom 50%. We first used cross-sectional data to examine expenditure concentration by calculating the proportion of all expenditures incurred by the top 10%, top 25% and top 50% of individuals in each year. Characteristics of individuals with AD in various high-expenditure groups (i.e., top 10%, top 25%, top 50% and bottom 50%) were examined. Student's t-tests and chi-square tests were performed to assess the statistical significance of differences in characteristics between the top 10%, top 25% and top 50% versus the lower 50% of individuals. A transition probability matrix was presented to illustrate the change in expenditure percentiles from one year to a subsequent year [Monheit, 2003; Liptak et al., 2006]. Ordered logit models [Maddala, 1983] were performed to predict the probability of being in a certain expenditure group in the next year, accounting for the ordinal nature of the

outcome variable (i.e., top 10%, next 11%-25%, next 26%-50% and bottom 50%). All analyses were conducted using SAS version 8.2 (SAS Institute, Cary, NC).

5.4 Results

Expenditure concentration

In 2007 constant dollars, average per capita health expenditures of elderly, community-dwelling Medicare beneficiaries with AD were $19,927 in 1999 and $22,144 in 2004, representing an 11% increase. The increase was more dramatic for prescription drug expenditures, from $1,797 in 1999 to $2,943 in 2004, a 64% increase, possibly due to increased use of prescription drugs, new indications for existing drugs, growth in therapeutic classes, and increased use of specialty drugs [CMS, 2007b]. As shown in Figures 5.2 and 5.3, overall and prescription drug expenditures were highly concentrated, with a relatively small proportion of high-cost users accounting for a large share of expenditures. The top 10% of individuals accounted for 38%-47% of overall health expenditures; the top 25% individuals accounted for nearly 70% and the top 50% accounted for nearly 90% of overall expenditures. Prescription drug expenditures appeared less concentrated than did overall health expenditures. Nevertheless, the top 10% made up approximately one-third, the top 25% spent nearly 60% and the top 50% incurred 85% of overall drug expenditures.

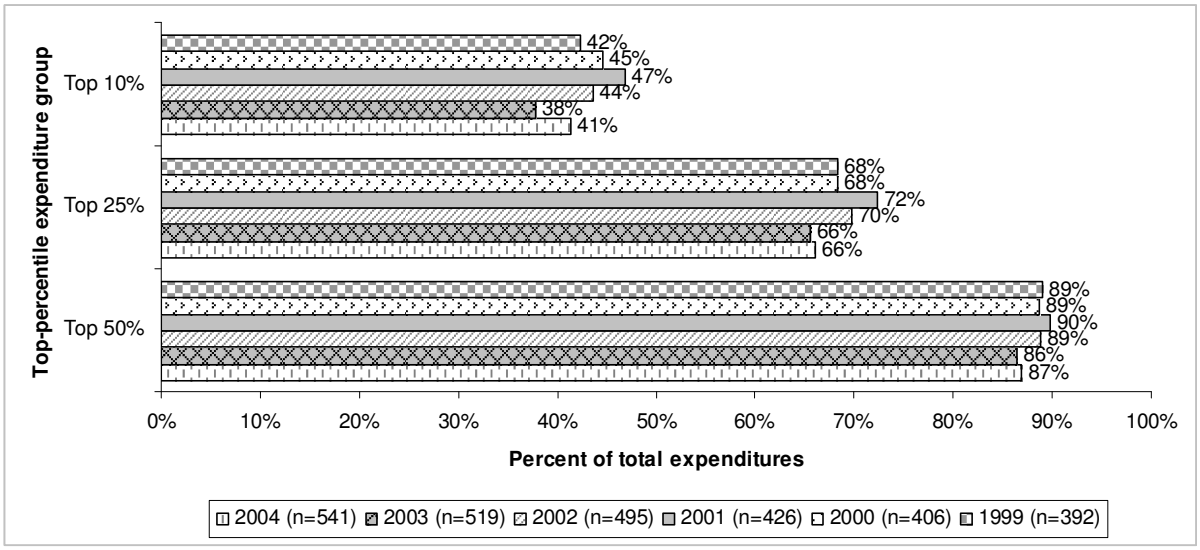


Figure 5.2: Concentration of health care expenditures in individuals with AD by top-percentile expenditure group and by year (1999-2004)

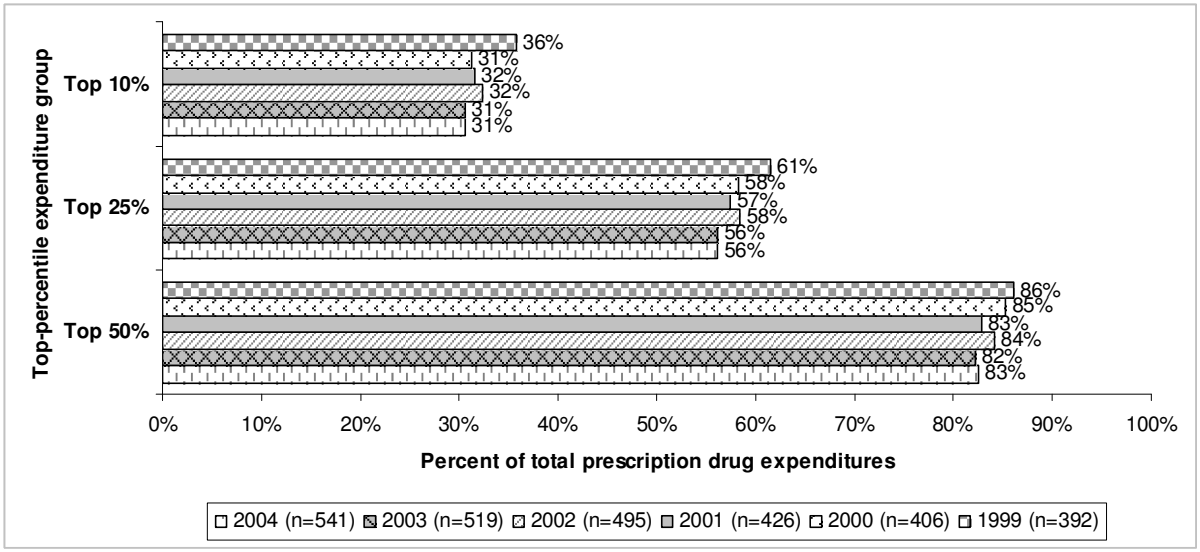


Figure 5.3: Concentration of prescription drug expenditures in individuals with AD by top-percentile expenditure group and by year (1999-2004)

Table 5.1 describes the demographic characteristics, health status and top five therapeutic classes for prescription medicines among persons with AD in various percentiles of total health expenditures. Individuals with AD who were ranked in the upper percentiles reported poorer health status relative to those in the bottom 50%. For instance, 72% of those in the top decile rated themselves to be in fair or poor health, in contrast with only 35% in the

bottom half. Comorbidities also were more pronounced in persons with AD in the upper percentiles, with an average of 5.4 (standard deviation [s.d.]=3.6) conditions in the top 10%, 4.8 (s.d.=3.4) in the top 25% and 4.1 (s.d.=3.1) in the top 50%, in contrast with an average of 1.9 (s.d.=2.0) conditions in the bottom half (all p-values<0.001).

Among individuals ranked in the top 10%, top 25% and top 50%, cardiac and cardiovascular medicines are the two most-frequently used therapeutic classes of prescription medications, followed by psychotherapeutic drugs, gastrointestinal preparations and autonomic agents (e.g., antihypertensive drugs). Individuals ranked in the bottom 50% also had the same drug classes ranked in the top five, whereas cardiovascular and autonomic medicines were the most frequently used.

Table 5.1: Characteristics of individuals with Alzheimer's disease by percentile of total health care expenditures (n=2,779)

Characteristics	Percentile group				p-value (Top 50% vs. Bottom 50%)
	Top 10% (n=281)	Top 25% (n=697)	Top 50% (n=1,391)	Bottom 50% (n=1,388)	
Mean age in years (s.d.)	81.3 (6.8)	82.0 (7.1)	82.2 (7.2)	81.7 (7.2)	0.095
Male, %	42.4	41.3	39.8	36.7	0.087
Race/ethnicity, %					0.518
White, not Hispanic	77.9	80.6	81.7	82.9	
Black, not Hispanic	16.4	13.4	12.6	11.6	
Hispanic	2.5	2.3	2.5	1.9	
Other	3.2	3.7	3.2	3.6	
Proxy respondent, %	64.4	64.0	58.2	46.5	<0.001
General health status, %					<0.001
Excellent / very good	9.3	15.5	18.2	33.2	
Good	18.3	21.0	13.0	31.7	
Fair	34.0	32.7	26.1	23.9	
Poor	38.4	30.8	31.2	11.2	
Charlson Comorbidity Index, mean (s.d.)	5.4 (3.6)	4.8 (3.4)	4.1 (3.1)	1.9 (2.0)	<0.001
Memory loss interfering with daily activities, %	55.8	58.5	59.8	58.6	0.536
Decision-making problems interfering with daily activities, %	51.6	53.1	52.4	49.9	0.182
Top 5 therapeutic classes for prescription medications, %					<0.001
Cardiac drugs	11.0	10.9	10.9	11.1	
Cardiovascular drugs	10.4	11.0	11.7	15.2	
Psychotherapeutic drugs	9.2	9.2	10.1	9.1	
Gastrointestinal preparations	7.7	7.5	7.1	5.4	
Autonomic drugs	6.6	6.9	7.5	11.8	

s.d.: standard deviation

Functional limitations were much more pronounced in the upper percentiles than in the bottom-half users (Table 5.2): the top 10% had an average of 2.4 (s.d.=2.3) ADL and 3.4 (s.d.=2.2) IADL impairments, compared with 1.0 (s.d.=1.8) ADL and 2.3 (s.d.=3.3) IADL limitations in the lower 50% of individuals (all p-values <0.001). Of the six ADL items assessed, bathing and dressing were the two most prevalent ADL impairments, and eating was the least prevalent. The Katz hierarchical index also suggested more ADL impairments among individuals with high expenditures. More than 23% of the highest-spending 10% of users, and more than 19% of the highest-spending 25% of users were impaired on five or all six ADLs (levels F and G), compared with 11.4% of the bottom 50% of users. Only 28% of the individuals in the top decile were fully independent on all ADLs (level A), compared with 53% of the bottom-half users. Performing IADLs is more cognitively demanding than ADLs, making them more prevalent in community-dwelling individuals with AD. Of the six IADLs assessed, individuals with AD were more likely to received help with or be unable to perform heavy housework due to health problems, but were less likely to have difficulty using the telephone.

Table 5.2: Functional status of individuals with Alzheimer's disease by percentile of total health care expenditures

Functional status	Percentile group				p-value (Top 50% vs. Bottom 50%)
	Top 10% (n=281)	Top 25% (n=697)	Top 50% (n=1,391)	Bottom 50% (n=1,388)	
Number of ADL impairments, mean (s.d.)	2.4 (2.3)	2.1 (2.2)	1.8 (2.2)	1.0 (1.8)	<0.001
ADL limitation, %					
Bathing	55.2	50.4	43.9	27.4	<0.001
Dressing	48.8	43.0	36.2	21.3	<0.001
Eating	19.4	18.9	15.2	7.4	<0.001
Transferring in and out of chairs	43.4	35.8	30.0	15.9	<0.001
Walking	38.0	32.6	28.2	16.6	<0.001
Toileting	30.8	26.6	22.5	12.2	<0.001
Katz index of ADLs, %					<0.001
A: independent in all activities	27.6	28.0	33.9	52.6	
B: independent in all but one activity	16.1	21.9	22.3	21.0	
C: dependent in bathing and one other activity	7.2	7.1	7.2	5.5	
D: dependent in bathing, dressing and one other activity	6.1	6.8	5.8	4.2	
E: dependent in bathing, dressing, going to toilet and one other activity	4.7	3.9	3.0	1.5	
F: dependent in bathing, dressing, going to the toilet, transferring and one other activity	12.5	9.1	8.3	3.8	
G: dependent in all activities	10.7	10.2	8.1	4.4	
Other: dependent in at least two activities and not classified in A-G	15.1	13.0	11.4	7.0	
Number of IADL impairments, mean (s.d.)	3.4 (2.2)	3.3 (2.3)	3.0 (2.4)	2.3 (3.3)	<0.001
IADL limitation, %					
Using the phone	31.9	31.9	29.0	21.9	<0.001
Doing light housework	60.6	54.7	49.3	31.3	<0.001
Doing heavy housework	73.5	69.7	65.7	48.5	<0.001
Making meals	56.3	53.5	48.8	37.0	<0.001
Shopping	64.2	60.3	56.1	44.1	<0.001
Managing money	57.4	56.6	53.3	46.3	<0.001

ADL: activity of daily living; IADL: instrumental activity of daily living; s.d.: standard deviation.

Expenditure persistence

Table 5.3 shows the change in the position in the total expenditure distribution and drug expenditure distribution during a two-year period. This matrix illustrates the likelihood of being in the various percentiles for the subsequent year, given the position in prior year's expenditure distribution (i.e., transition probability). For example, an individual in the top 10% of the distribution would have a probability of 0.25 for remaining in the top decile and a probability of 0.21 for moving to the bottom half in the next year. Individuals in the bottom-half

of the expenditure distribution would have a probability of only 0.05 for moving up to the top decile but 0.70 for retaining this ranking in the next year. Nearly 70% of those in the bottom 50% stayed in the bottom half for two consecutive years.

Prescription drug expenditures exhibited even more persistence than did overall health expenditures. The highest-spending 10% of users would have a probability of 0.50 for retaining this ranking, 0.30 for being in the next 11%-25%, and only 0.09 for becoming the bottom half. Individuals in the bottom 50% of the drug expenditure distribution would have a probability of only 0.01 for moving up to the top decile, and a probability of only 0.03 for moving to the next highest-spending 11%-25%, whereas 75% of those in the bottom 50% retained the this ranking in the next year. This matrix confirmed that a large proportion of individuals with AD who were in the top percentiles of overall and prescription drug expenditures retained this ranking in a second year, whereas most individuals with AD in the bottom half of the expenditure distribution remained in the low-spending group in the subsequent year.

Table 5.3: Transition probability matrix for expenditure percentiles moving from the baseline year to the subsequent year (1999-2004 pooled data, n=671)

Position in base-year expenditure distribution	Position in subsequent-year expenditure distribution			
	Top 10%	Next 11%-25%	Next 26%-50%	Bottom 50%
	Total health care expenditures			
Top 10%	0.25	0.17	0.37	0.21
Next 11%-25%	0.19	0.26	0.24	0.31
Next 26%-50%	0.08	0.16	0.35	0.41
Bottom 50%	0.05	0.09	0.19	0.70
	Prescription drug expenditures			
Top 10%	0.50	0.30	0.11	0.09
Next 11%-25%	0.18	0.33	0.28	0.21
Next 26%-50%	0.06	0.20	0.43	0.31
Bottom 50%	0.01	0.03	0.21	0.75

Predicting the likelihood of future expenditure level

Prior expenditures had substantial predictive power, much stronger than demographic characteristics and functional health status, in predicting the likelihood of being in a certain expenditure group in the next year (Table 5.4). An individual's position in the

base-year total expenditure distribution was a strong predictor of high expenditures in a second year. Individuals in the top 10%, next 11%-25% and next 26%-50% all were much more likely to retain this ranking compared with those in the bottom 50% ($p<0.001$). Based on the model with simple counts of ADL and IADL impairments, we calculated the marginal effects of prior expenditure category (Table 5.5). The highest-spending 10% of users were 16.3-percentage points more likely to be in the top decile in the next year, compared with the bottom-half of users in the expenditure distribution. The top 10% of users were 12.7-percentage points more likely to become the next 11%-25% in the subsequent-year expenditure distribution, and were 32.5-percentage points less likely to become the bottom 50%. The likelihood of being in the upper percentiles also increased with the number of comorbidities measured by the CCI ($p<0.001$). Increase in one comorbidity category would result in 2.0-percentage point increase in the likelihood of being in the top 10% in the expenditure distribution and 5.8-percentage point decrease in the likelihood of being the bottom 50% in the subsequent year.

Table 5.4: Ordered logit models for the likelihood of persistently high health care expenditures: Predicting the position in subsequent-year expenditure distribution (1999-2004 pooled data, n=671)

Characteristics in the baseline year	Model 1 Coefficient (s.e.)	Model 2 Coefficient (s.e.)	Model 3 Coefficient (s.e.)
Position in base-year total expenditure distribution			
Top 10%	1.69 (0.31)***	1.79 (0.32)***	1.79 (0.32)***
Next 11%-25%	1.44 (0.24)***	1.38 (0.24)***	1.38 (0.24)***
Next 26%-50%	0.83 (0.19)***	0.79 (0.20)***	0.78 (0.20)***
Bottom 50%	--[1]	--	--
Age	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Male sex	-0.10 (0.17)	0.06 (0.17)	0.11 (0.17)
Race/ethnicity			
White, not Hispanic	--	--	--
Black, not Hispanic	0.07 (0.25)	0.09 (0.25)	0.08 (0.25)
Hispanic or other	0.10 (0.34)	- 0.02 (0.35)	0.16 (0.35)
Charlson Comorbidity Index	0.29 (0.03)***	0.29 (0.03)***	0.29 (0.03)***
General Health status			
Excellent/very good	--	--	--
Good	0.04 (0.22)	-0.03 (0.22)	-0.08 (0.22)
Fair	-0.10 (0.23)	-0.20 (0.24)	-0.27 (0.24)
Poor	0.32 (0.27)	0.18 (0.28)	0.13 (0.28)
Memory loss interfering with daily activities	-0.08 (0.22)	-0.10 (0.22)	-0.16 (0.22)
Decision-making problems interfering with daily activities	-0.12 (0.24)	0.11 (0.24)	0.09 (0.24)
Functional status			
Number of ADL impairments	0.10 (0.06)*		
Number of IADL impairments	-0.03 (0.05)		
Dummy variables of ADL impairments			
Bathing		0.19 (0.28)	
Dressing		-0.06 (0.31)	
Eating		-0.24 (0.31)	
Transferring in and out of chairs		0.51 (0.30)*	
Walking		0.26 (0.27)	
Toileting		-0.04 (0.31)	
Dummy variables of IADL impairments			
Using the phone		-0.48 (0.23)**	-0.48 (0.22)**
Doing light housework		-0.02 (0.28)	0.0004 (0.28)
Doing heavy housework		0.70 (0.23)***	0.70 (0.23)***
Making meals		-0.22 (0.31)	-0.19 (0.31)
Shopping		0.26 (0.30)	0.20 (0.30)
Managing money		-0.65 (0.26)**	-0.62 (0.27)**
Katz index of ADL impairments[2]			
Level A			--
Level B			0.40 (0.23)*
Level C			0.55 (0.37)
Level D			0.54 (0.56)
Level E			-0.07 (0.54)
Level F			1.20 (0.38)***
Level G			0.50 (0.43)
Other			0.82 (0.34)**
Intercept 4	-5.25 (1.00)***	-5.16 (1.02)***	-5.20 (1.03)***
Intercept 3	-3.83 (0.98)***	-3.71 (1.01)***	-3.75 (1.01)***
Intercept 2	-2.37 (0.98)***	-2.21 (1.00)***	-2.23 (1.01)***
Pseudo R^2	0.257	0.285	0.290

[1] Reference group
[2] See Table 5.2 for Katz index classifications
ADL: activity of daily living; IADL: instrumental activity of daily living; s.e.: standard error.
Differences are statistically significant at *(p<0.10), **(p<0.05) and ***(p<0.01) levels

Table 5.5: Marginal effects of base-year total expenditures and comorbidities on subsequent-year total expenditures

Characteristics in the base year	Position in the subsequent-year total expenditure distribution			
	Top 10%	Next 11%-25%	Next 26%-50%	Bottom 50%
Position in the total expenditure distribution				
Top 10%	0.163	0.127	0.035	-0.325
Next 11%-25%	0.122	0.113	0.052	-0.288
Next 26%-50%	0.057	0.061	0.049	-0.167
Bottom 50%	--	--	--	--
Charlson Comorbidity Index	0.020	0.022	0.016	-0.058

Note: Predictions are based on the ordered logit model with simple numbers of functional limitations in Table 5.4 (Model 1)

Higher prior drug expenditures and higher CCI also significantly increased the likelihood of being in the upper percentiles for drug expenditures in the subsequent year, regardless of the form of functional status used in the ordered logit model (Table 5.6). Individuals with the highest (i.e., top 10%) drug expenditures were 56.5-percentage points more likely to retain this ranking, and were 52.0 percentage points less likely to move to the bottom-half in the expenditure distribution in the next year, compared to those with the lowest drug expenditures (i.e., bottom 50%) (Table 5.7). An increase in one comorbidity category in the CCI would result in a 0.5-percentage point increase in the likelihood of being in the top 10% in the subsequent year and a 1.1-percentage point decrease in the likelihood of being the bottom 50%.

Table 5.6: Ordered logit models for the likelihood of persistently high prescription drug expenditures: Predicting the position in subsequent-year expenditure distribution (n=671)

Characteristics in the baseline year	Model 1 Coefficient (s.e.)	Model 2 Coefficient (s.e.)	Model 3 Coefficient (s.e.)
Position in base-year drug expenditure distribution			
Top 10%	4.28 (0.30)***	4.32 (0.30)***	4.33 (0.30)***
Next 11%-25%	2.79 (0.24)***	2.85 (0.25)***	2.82 (0.25)***
Next 26%-50%	1.87 (0.20)***	1.96 (0.21)***	1.89 (0.20)***
Bottom 50%	--[1]	--	--
Age	-0.04 (0.01)***	-0.04 (0.01)***	-0.04 (0.01)***
Male	-0.28 (0.17)*	-0.26 (0.17)	0.24 (0.17)
Race/ethnicity			
White, not Hispanic	--	--	--
Black, not Hispanic	-0.89 (0.28)***	-0.84 (0.29)***	-0.91 (0.29)***
Hispanic or other	0.02 (0.35)	-0.07 (0.35)	-0.01 (0.35)
Charlson Comorbidity Index	0.07 (0.03)**	0.07 (0.03)**	0.06 (0.03)**
General health status			
Excellent/very good	--	--	--
Good	0.21 (0.21)	0.23 (0.22)	0.21 (0.22)
Fair	-0.01 (0.24)	0.04 (0.24)	-0.01 (0.24)
Poor	0.07 (0.28)	-0.02 (0.29)	-0.02 (0.28)
Memory loss interfering with daily activities	-0.27 (0.22)	-0.26 (0.22)	-0.29 (0.22)
Decision-making problems interfering with daily activities	-0.10 (0.24)	-0.06 (0.25)	-0.09 (0.25)
Functional status			
Number of ADL impairments	-0.02 (0.06)		
Number of IADL impairments	-0.02 (0.05)		
Dummy variables of ADL impairments			
Bathing		-0.67 (0.30)**	
Dressing		0.44 (0.32)	
Eating		-0.64 (0.34)	
Transferring in and out of chairs		-0.07 (0.33)	
Walking		0.56 (0.28)*	
Toileting		0.02 (0.33)	
Dummy variables of IADL impairments			
Using the phone		-0.38 (0.24)	-0.41 (0.24)*
Doing light housework		0.24 (0.29)	0.16 (0.29)
Doing heavy housework		-0.15 (0.23)	-0.11 (0.23)
Making meals		-0.12 (0.31)	-0.12 (0.32)
Shopping		0.41 (0.30)	0.43 (0.30)
Managing money		-0.15 (0.26)	-0.15 (0.26)
Katz index of ADL impairments[2]			
Level A			--
Level B			0.0006 (0.22)
Level C			-0.50 (0.39)
Level D			0.03 (0.60)
Level E			-0.55 (0.58)
Level F			0.41 (0.39)
Level G			-0.25 (0.45)
Other			0.01 (0.34)
Intercept 4	-1.07 (0.98)	-1.32 (1.01)	-1.11 (1.01)
Intercept 3	0.52 (0.98)	0.31 (1.01)	0.51 (1.01)
Intercept 2	2.29 (0.99)**	2.12 (1.01)**	2.30 (1.01)**
Pseudo R-squared	0.409	0.426	0.418

[1] Reference group
[2] See Table 5.2 for Katz index classifications
ADL: activity of daily living; IADL: instrumental activity of daily living; s.e.: standard error.
Differences are statistically significant at *(p<0.10), **(p<0.05) and ***(p<0.01) levels

Table 5.7: Marginal effects of base-year drug expenditures and comorbidities on subsequent-year drug expenditures

Characteristics in the base year	Position in the subsequent-year drug expenditure distribution			
	Top 10%	Next 11%-25%	Next 26%-50%	Bottom 50%
Position in the drug expenditure distribution				
Top 10%	0.565	0.114	-0.159	-0.520
Next 11%-25%	0.263	0.160	0.021	-0.443
Next 26%-50%	0.146	0.084	0.095	-0.325
Bottom 50% (reference group)	--	--	--	--
Charlson Comorbidity Index	0.005	0.004	0.002	-0.011

Note: Predictions are based on the ordered logit model with simple numbers of functional limitations in Table 5.6 (Model 1)

5.5 Discussion

This study examined the extent of overall and prescription drug expenditure concentration and persistence among Medicare beneficiaries with AD, a disease that places substantial burdens on both the individuals themselves and the health care system. Our data indicated that the highest-spending 10% of users accounted for as much as 47% of overall health expenditures and one-third of prescription drug expenditures among individuals with AD. At the other end of the spectrum, the lower half was responsible for only 10% of overall health expenditures and approximately 15% of drug expenditures. Compared with individuals in the bottom 50%, the upper percentiles for total expenditures had more comorbidities, poorer health status and more functional impairments.

Evidence has shown that expenditures for individuals with AD varied considerably by comorbid illness [Hill et al., 2002], and comorbid conditions were highly prevalent in our sample, especially among higher-expenditure individuals. The CCI scores significantly predicted future levels of expenditures. Given the fact that persons with AD are prone to forgetting what they were told to manage their comorbid illnesses or even losing their self-care skills, better treatments and disease management of AD are needed to facilitate care coordination and to reduce the costs of comorbidities commonly experienced by these frail elderly. These intervention programs may include use of appropriate medications (e.g., acetylcholinesterase inhibitors), care coordination by nurse case managers and/or primary

care physicians (e.g., discharge planning), education, counseling and support of caregivers [Fillit et al., 2002b; Kaufer et al., 2005].

Prescription drug expenditures appeared less concentrated than total expenditures probably because infrequent and expensive events such as extended hospitalizations make the distribution of overall expenditures more right-skewed [Banthin and Miller, 2006]. Nevertheless, drug expenditures exhibited high persistence in this AD population—half of the highest 10% retained this ranking but only 9% dropped to the bottom 50% in the expenditure distribution in a second year. Persistence in health expenditures may be distributed randomly or may be correlated with personal characteristics or previous expenditures and, therefore, can be predicted. Our analysis demonstrated that an individual's position in the prior-year expenditure distribution significantly predicted the expenditure percentiles in the subsequent year. The greater the degree of high-expenditure persistence, the more likely it is that adverse selection into Medicare prescription drug plans will occur [Banthin and Miller, 2006]. Under the new Medicare Part D, beneficiaries may self-select their prescription drug coverage. Whether beneficiaries decide to enroll may be based on weighing the cost of enrollment against their anticipated need for prescription drugs in the future [Coulson and Stuart, 1992]. Our data showed that, among individuals with AD, cardiac, cardiovascular, psychotherapeutic, gastrointestinal and autonomic drugs were the five most commonly used classes of prescription medicines. These therapeutic classes are different from those used by general community-dwelling Medicare beneficiaries who use more diuretics and antiarthritics in addition to cardiac, cardiovascular and gastrointestinal medicines [Simoni-Wastila et al., 2007]. Our findings help policymakers to design formularies that offer better and more comprehensive drug coverage for seniors with AD to meet their different needs. Identifying the characteristics of individuals with high and persistently-high expenditures may help to develop equitable health insurance strategies, such as catastrophic care, carve-outs, reinsurance and risk adjustment [Liptak et al., 2006], especially in a situation where medical

services for the long-term illnesses are expected to cost more than insurance coverage limits or more than the amount most families expect to pay with their own resources.

An important issue that affects the dynamics of expenditures is regression to the mean, a statistical phenomenon in which the lower-spending cohorts increase and the highest-spending cohorts decrease their expenditures (towards the group mean) over time [Beebe, 1988; Barnett et al., 2005]. Our data revealed that regression to the mean occurred around both total expenditures and drug expenditures during a two-year period. For instance, of those top decile (bottom half) with the highest (lowest) health expenditures, 37% decreased (19% increased) their expenditures and moved to the middle-expenditure group (i.e., next 26%-50% column in Table 5.3) in a second year. Of those top 10% (bottom 50%) individuals with the highest (lowest) drug expenditures, 11% decreased (21% increased) their expenditures and became the middle-expenditure group. However, we were unable to determine to what degree the change in expenditure percentiles was due to improvement in outcomes of care, change in disease severity or simple regression to the mean. Coulson and Stuart [1992] examined the persistence in pharmaceutical use by the elderly and found that regression to the mean may take more than four years to complete. In the short run, the two highest-spending cohorts were found to be well above the overall mean, and the three lowest cohorts were still well below the group mean. We did not have a sufficient number of years of data to investigate regression to the mean beyond two years. In the future, longer follow-up periods may be necessary to accurately quantify this phenomenon.

Previous research using cross-sectional data suggests a strong relationship between functional impairment and health expenditures [Hill et al., 2006]. However, most of the functional limitations we tested did not significantly predict future expenditure percentiles. We suspect that the lack of significance was due to multicollinearity based on the findings that ADLs and IADLs were strongly correlated (Pearson correlation coefficient $r=0.68$, $p<0.001$), and that both functional disability measures were weakly correlated with prior expenditure

percentiles ($r=0.12$-0.29, $p<0.001$). Note that functional limitations were highly prevalent in our sample, especially among those in the high-expenditure percentiles, indicating that these individuals may have greater non-health care needs, such as requiring assistance with housework and managing money, in addition to medical attention.

Several limitations merit discussion. Our sample excluded individuals with any managed care participation (n=419) because their medical claims were not available in the MCBS. In general, these individuals were somewhat younger and healthier and had lower average overall and drug expenditures than our study sample. To provide a point of comparison, we also examined expenditure concentration and persistence including MCO enrollees and the results were almost numerically identical with our findings for fee-for-service members. We also substituted CCI scores with patient- or proxy-reported conditions in the ordered logit models. These conditions included heart disease, cancer, diabetes, stroke, high blood pressure, chronic obstruction pulmonary disease, paralysis, arthritis and mental illness. Prior expenditures remained a strong predictor of the level of future overall and drug expenditures, whereas only high blood pressure, but not other comorbid conditions, was statistically significant. Finally, our sample excluded individuals residing in skilled nursing facilities who generally have more functional impairments and more comorbidities than community-dwelling individuals with AD [Hill et al., 2006]. The scope of future studies should be expanded to include data from large Medicare MCOs and long-term-care facilities that might cover persons with AD with different utilization patterns and expenditures.

5.6 Conclusion

Our analysis of expenditure data from a nationally representative sample of Medicare beneficiaries with AD revealed that both overall and prescription drug expenditures were highly concentrated in a small group of high-expenditure individuals and were persistent over

a two-year period. Prescription drug expenditures exhibited less concentration but more persistence than did overall health expenditures in this elderly AD population. Our results highlight the importance of prior expenditures and comorbidities in predicting the level of future expenditures.

CHAPTER 6

RISK ADJUSTMENT FOR MEDICARE BENEFICIARIES WITH ALZHEIMER'S DISEASE

6.1 Abstract

Introduction: Previous research has raised a concern about the adequacy and accuracy of risk adjustment models based solely on diagnosis codes in administrative data to make capitation payments for health plans with mostly frail or disabled populations. The objective of this research was to compare the performance of various prospective risk adjustment measures to predict overall and prescription drug expenditures of Medicare beneficiaries with Alzheimer's disease (AD). We also compared the measures in predicting being in the top 10% of the expenditure distribution.

Methods: Elderly, community-dwelling individuals with AD were identified from the 1999-2004 Medicare Current Beneficiary Survey linked with Medicare claims. Risk adjustment measures evaluated for this sample were constructed from diagnoses in medical claims, functional status in survey report, and prior expenditures. Model performance in expenditure analysis was compared using adjusted R^2, log likelihood values, predictive ratios, and receiver operating characteristic (ROC) curves in high-expenditure outlier analysis.

Results: Adding diagnosis-based measures to the demographic model substantially increased the adjusted R^2 from 1% to 17.0% with the Charlson Comorbidity Index (CCI), followed by 15.5% with the Chronic Illness and Disability Payment System-Medicare (CDPSM), and 12.1% with the Centers for Medicare and Medicaid Services modified version of Hierarchical Condition Category (CMS-HCC) model. Prior expenditures were more predictive (adjusted R^2=20.4%) than any other single measures. Incorporating the frailty

adjuster based on limitations of activity of daily living modestly increased the adjusted R^2 of the diagnosis-based models, with the CCI remaining the most predictive (18.5%), but not the prior-expenditure model. All three diagnosis-based measures under-predicted the actual total expenditures in the highest quartile by approximately 50%, and over-predicted those in the lowest quartile by six fold. For drug expenditures, all three diagnosis-based measures performed similarly and produced greater over-prediction in the lowest quartile than for total expenditures. Prior drug expenditures were more accurate than any other measures but tended to over-predict across quartiles. Given a 10% high-expenditure threshold, the area under the ROC curves measured by the *c*-statistics were highest in the CCI (0.667), followed by the CDPSM (0.645) and CMS-HCC (0.628). For drug expenditures, no model except past use predicted high-expenditure outliers well.

Conclusions: In single-measure, diagnosis-based models, the CCI outperformed the CMS-HCC and the CDPSM in predicting total and drug expenditures in our sample with AD. A frailty adjuster based on ADL limitations improved overall prediction and predictive accuracy of the diagnosis-based models, especially the CMS-HCC specification currently used by Medicare. Only prior expenditures, but not diagnoses, predicted drug expenditures in the next year with any accuracy, and can discriminate between high- and low-expenditure individuals. Future research is needed to evaluate the performance of risk adjustment measures based on ambulatory pharmacy data.

Keywords: Alzheimer's disease, risk adjustment, Medicare, health expenditures, drug expenditures

6.6 Introduction

Predictive modeling techniques increasingly have been used to forecast health care services use and expenditures in managed care [Pope et al., 2000; Powers et al., 2005; Zhao

et al., 2005]. One of the most important applications of predictive models is for making risk-adjusted payments. The Balanced Budget Act of 1997 required the Centers for Medicare and Medicaid (CMS, formerly Health Care Financing Administration) to implement risk-adjusted Medicare capitation payments for managed care plans by January 1, 2000, in order to compensate fairly health plans for the expected expenditures associated with the disease burden of their enrollees. CMS first adopted a form of risk adjustment based on inpatient hospital diagnoses (i.e., using the principal inpatient diagnostic cost groups system, PIPDCG) [Pope et al., 2000]. Beginning 2004, CMS applied Medicare-specific modifications to the Diagnostic Cost Group-Hierarchical Condition Category (DCG-HCC) model, which incorporates both inpatient and outpatient claims, as the new risk adjustment capitation payment system (hereafter known as CMS-HCC).

Risk adjustment models generally rely on some combination of individual-level characteristics to predict outcomes, such as mortality or health expenditures. Studies of prospective models of expenditures have examined various risk adjustment measures, such as diagnoses [Ellis et al., 1996; Meenan et al., 1999; Ash et al., 2000; Pope et al., 2000; Riley, 2000; Meenan et al., 2003; Hughes et al., 2004; Pope et al., 2004; Noyes et al., 2006], pharmacy claims data [Von Korff et al., 1992; Fishman et al., 2003; Powers et al., 2005], prior utilization [Ash et al., 2001; Monheit, 2003], health status from survey data [Epstein and Cumella, 1988; Fowles et al., 1996; Pope et al., 1998; Lamers, 1999; Temkin-Greener et al., 2001; Pacala et al., 2003; Kautter and Pope, 2004; Fleishman et al., 2006], or some combination thereof [Zhao et al., 2001; Maciejewski et al., 2005; Zhao et al., 2005; Farley et al., 2006]. These studies generally find that combined models, such as diagnoses plus other risk adjusters, explain more variation in expenditures and improve predictive power, relative to single-measure models (e.g., diagnoses alone), probably because different risk adjustment measures explain different segments of the distribution.

The performance of a risk adjustment model can vary across different demographic groups and disease populations. For instance, compared with other diagnosis- or pharmacy-based measures, the DCG-HCC model has been shown to capture the largest proportion of total expenditures among MCO beneficiaries with asthma, diabetes, or depression [Meenan et al., 2003]. However, observers report that diagnosis-based risk adjustment measures do not adequately compensate health plans serving primarily disabled or frail populations, defined in terms of functional impairment [Riley, 2000; Robinson and Karon, 2000; Temkin-Greener et al., 2001; Kautter and Pope, 2004]. Kautter and Pope [2004] found that, without frailty adjustment, the CMS-HCC model would under-predict Medicare expenditures by an average of $4,923, $1,531, and $809, respectively, for beneficiaries with 5-6, 3-4, and 1-2 impairments of activities of daily living (ADLs, i.e., bathing, dressing, eating, transferring in and out of chairs, walking, and toileting), and over-predict by $697 for those with no ADL difficulties. As risk adjustment models continue to be developed, further refinement is necessary for certain Medicare subpopulations, such as individuals with functional limitations.

This study compared the performance of various generic risk adjustment models in Medicare beneficiaries with Alzheimer's disease (AD), a population with progressive decline in cognition, memory and functional ability. It is estimated that more than 60% of individuals with AD have three or more comorbid conditions [Doraiswamy et al., 2002], and that comorbidities are an important driver of increased costs for AD [Gutterman et al., 1999]. Particularly, ADL limitations are a critical part of the disease progression [Mohs et al., 2000] which are correlated strongly with health care expenditures [Taylor et al., 2001; Leung et al., 2003; Fillit and Hill, 2004; Hill et al., 2006]. Medicare pays for most of the cost of hospitalization and a large portion of other medical care among elderly individuals with AD. In 2000, individuals with AD represented fewer than 5% of Medicare beneficiaries [Taylor and Sloan, 2000], whereas they accounted for 14.4% of overall Medicare spending [Alzheimer's

Association, 2001]. Despite its progressive disabilities and high cost, AD is not accounted for in the current CMS-HCC diagnostic classification system. As health care spending for Medicare beneficiaries continues to rise, there is a critical need to identify prospectively high-expenditure groups, predict resource use, and control expenditures using tools such as risk adjustment models.

The objective of this study was to evaluate the predictive accuracy of five risk adjustment measures, including three diagnosis-based risk adjustment measures derived from medical claims records, one frailty adjuster based on functional status (i.e., ADL limitations) collected from survey data, and prior expenditures, in prospective models of overall and prescription drug expenditures. This study also examined the performance of models combining the frailty adjuster and other measures of risk. This comparison sheds light on refining current risk-adjusted payment methods, and helps plan managers who seek productive areas for disease management investment by identifying future high-expenditure cases [Leon et al., 1998; Fillit, 2000].

6.6 Methods

Data Source

Data were obtained from the 1999-2004 waves of the Medicare Current Beneficiary Survey (MCBS) Cost and Use files, linked with Medicare claims data. The MCBS is a nationally representative and comprehensive survey of health care use, expenditures, sources of payment and health status for the Medicare population [CMS, 2006b]. It provides a unique opportunity to study risk adjustment models because the data integrate information directly from a beneficiary with his/her Medicare claims for provider services and covered charges [CMS, 2006b]. Available Medicare Part A (i.e., hospitalization, skilled nursing facility, hospice, and home health care) and Part B (i.e., physician visits, specified outpatient care and some outpatient medications) claims records include diagnosis codes, utilization,

charges and reimbursement for all medical services [CMS, 2003]. The survey oversamples significant subpopulations (e.g., older beneficiaries) allowing increased statistical power for subgroup analyses, such as for elderly persons with AD. Survey interviews are completed by either sample members or proxy respondents (usually a family member or close acquaintance) if the sample member is unable to respond due to physical or mental problems. Each individual is interviewed three times a year to form a continuous profile of the individual's health care experience [CMS, 2006b]. The survey features a longitudinal rotating panel design, in which sampled individuals remain in the panel for no more than four years by the time they are retired, resulting in a cumulative sample size of three 4,000-beneficiary cohorts or approximately 12,000 beneficiaries in any given year. Therefore, the data set allows us to construct risk adjustment measures using base-year information to predict subsequent-year expenditures.

Sample

The study sample consisted of elderly, community-dwelling beneficiaries, defined as adults aged 65 and older who were not institutionalized for more than 90 days at a time during a year [CMS, 2006c] (Figure 6.1). To identify individuals with AD, we incorporated three case definitions because use of a single definition, such as an AD diagnosis in medical claims data, to define Alzheimer's cases has been shown to introduce errors of omission and commission [Newcomer et al., 1999; Fillit, 2000; Rice et al., 2001; Pressley et al., 2003; Lin, 2008a]. Therefore, we defined individuals who met any of the following case definitions as having AD:

1) affirmative answer to the question "Has a doctor ever told you that you had Alzheimer's disease or dementia?"; or

2) at least one AD diagnosis, defined by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes: all 290 codes (senile and presenile organic psychotic conditions) and 331.0 (AD) in Medicare Part A or Part B claims files; or

3) use of any Alzheimer's medications, including donepezil [Aricept®], rivastigmine [Exelon®], galantamine [Reminyl® or Razadyne®]) and memantine (Namenda®). These medications were identified by the drug names in survey-reported data, i.e., pharmacy administration data were not used.

Individuals who had any managed care participation during that year were excluded because these managed care organizations (MCOs) do not submit claims with diagnoses to the CMS, which we require to construct diagnosis-based risk adjustment measures. Of the 1,861 eligible, unique individuals, we excluded persons with only one year of claims (including deaths of sample respondents and nonresponse in later rounds). We then retained individuals who had medical claims data for both year *t* and year *t+1*, and excluded data for year *t+2* among those who were observed more than twice in the data set. As a result, the study sample consists of two-year panels of 671 unique individuals with AD. Of these, 155 (23%) were in the 1999-2000 cohort, 102 (15%) were in the 2000-2001 cohort, 136 (20%) were in the 2001-2002 cohort, 147 (22%) were in the 2002-2003 cohort, and 131 (20%) were in the 2003-2004 cohort; 79 (11.8%) died during the second year.

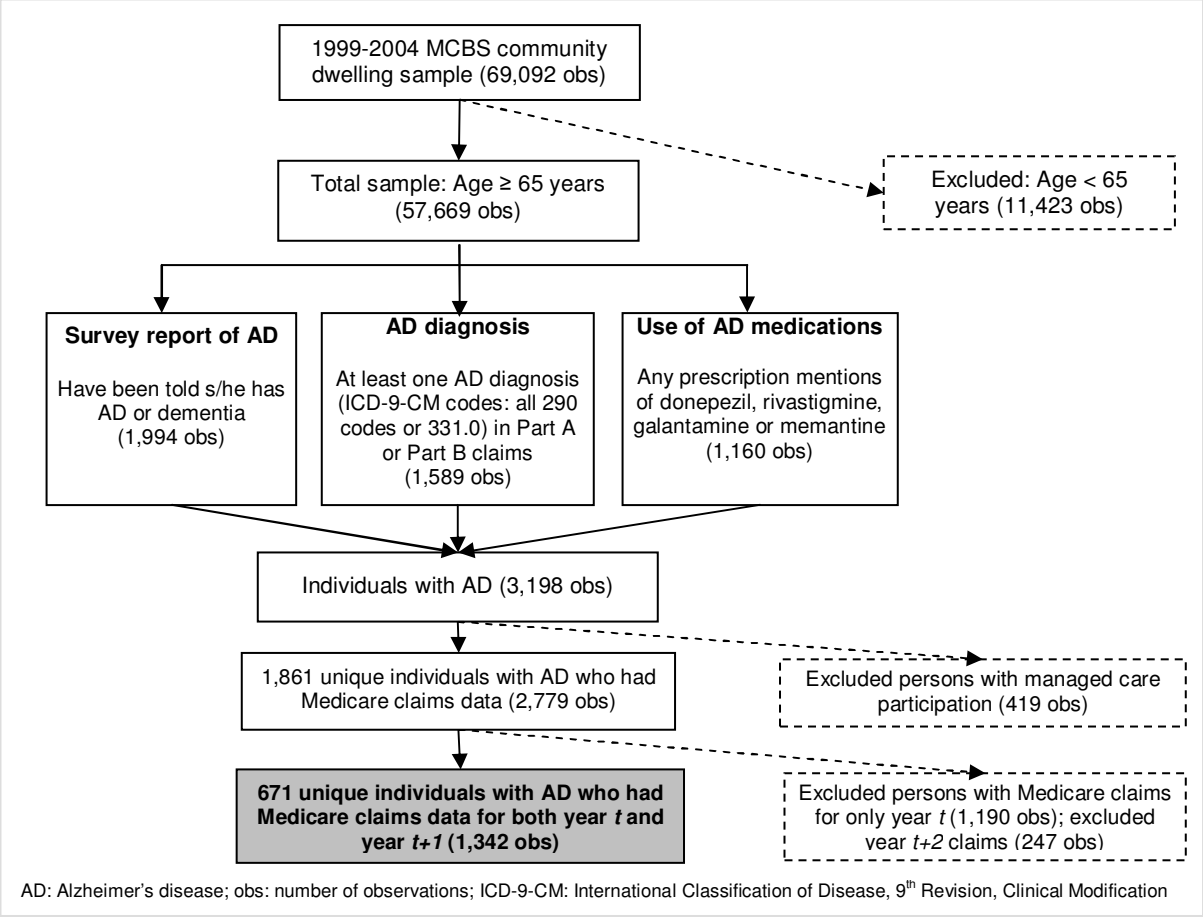


Figure 6.1: Sample extraction flowchart

Measures

Overall and prescription drug expenditures

Expenditure data in the MCBS were consolidated from both survey information and Medicare administrative files [CMS, 2003]. Personal total health care expenditures are defined as the MCBS aggregated payments across all types of services, not including rebates that may be paid by the manufacturer to the insurer. Prescription drug expenditures, a component of total expenditures, are imputed payments in the MCBS for prescription medicines received from all sources, including Medicare, Medicaid, Medicare MCOs, private MCOs, the Veterans Administration, employer-sponsored private insurance, individually-purchased private insurance, private insurance from an unknown source, out-of-pocket payments and public health plans other than Medicare or Medicaid [CMS,

2003]. Detailed procedures and criteria used in the MCBS to impute missing payments for medical services and prescription drugs are discussed elsewhere [England et al., 1994]. Briefly, a full set of internally consistent expenditure and payment records were created from these sources with minimal partial data discarded. We converted all expenditures into constant 2007 dollars using the Consumer Price Index for medical care [Bureau of Labor Statistics, 2007], which reflects the average price paid for a typical market basket of health care services.

Risk adjustment measures

The simplest risk adjustment measure includes only age categories (e.g., 65-69, 70-74, 75-79, 80-84, and 85) and gender. Other risk adjusters are discussed individually below.

CMS Diagnostic Cost Group-Hierarchical Condition Category (CMS-HCC)

Individuals are assigned to multiple Hierarchical Condition Categories (HCCs) based on demographic characteristics and diagnoses in both inpatient and outpatient claims in the prior year. The model also incorporates original reason for Medicare entitlement (disability or age), location of residence (community or facility), recent enrollment (covered by Medicare for fewer than 12 months in the prior year), and Medicaid eligibility. The model aggregates clinically and economically meaningful disease categories into 70 HCCs (e.g., metastatic cancer and acute leukemia) to which payment weights are assigned to reflect the level of increased expenditures associated with the HCC. We employed the HCC scores, which is the sum of the weights for all HCCs, of each individual as a risk adjustment measure. The CMS-HCC algorithm was obtained from the CMS website [CMS, 2006a].

Chronic Illness and Disability Payment System-Medicare (CDPSM)

The Chronic Illness and Disability Payment System (CDPS) is an expansion of a prior model, the Disability Payment System, developed by Kronick and colleagues [Kronick et al.,

1996]. At least eight state Medicaid programs use this diagnostic classification system to make risk-adjusted capitation payments for beneficiaries with disabilities or who receive Temporary Assistance to Needy Families [Kronick et al., 2000]. We used the CDPS Medicare version (CDPSM), a modified model more appropriate to predict expenditures for Medicare beneficiaries [Kronick et al., 2002]. Compared with the CMS-HCC, the CDPSM is more conservative in counting diagnoses and, therefore, has fewer categories. The CDPSM uses ICD-9-CM codes in both inpatient and outpatient claims to create 16 major disease categories, which correspond to body systems or type of diagnosis. The major categories are divided further into 66 subcategories. Similarly to the CMS-HCC, payment weights are assigned to those subcategories to reflect the level of increased expenditures associated with the condition. The CDPSM score then sums the weights for all indicated subcategories. The CDPSM software was acquired under license agreement at no charge (http://cdps.ucsd.edu, last accessed on 25 January, 2008).

Charlson Comorbidity Index (CCI)

The original CCI is a general comorbidity index consisting of 19 disease categories developed to predict one-year mortality using hospital chart review data [Charlson et al., 1987]. Other researchers have adapted the CCI for use with comorbidity data from administrative databases that include ICD-9-CM diagnosis codes [Deyo et al., 1992; D'Hoore et al., 1993; Romano et al., 1993; Ghali et al., 1996; Quan et al., 2002]. We used the CCI with the Deyo modification containing 17 comorbidity categories [Charlson et al., 1987; Deyo et al., 1992]. Each condition is assigned a weight as 1, 2, 3, or 6, reflecting the magnitude of the adjusted relative risks associated with each comorbidity. The CCI score then sums the weights for all conditions, with higher numbers representing a greater burden of comorbidity. The CCI for this study was derived using the presence of various ICD-9-CM codes in Medicare claims at the base year.

Functional status

Kautter and Pope [2004] developed a CMS frailty adjustment model that incorporates functional status in addition to the CMS-HCC to adjust for capitation payments to certain health plans specializing in providing care to the community-dwelling, frail elderly. The use of functional ability frailty adjuster, in particular measured by ADLs, is preferred to other health status measures (such as general perceived health status "Is your health excellent, very good, good, fair, or poor?") because it has good face validity and has been shown to explain Medicare expenditures not accounted for by diagnosis-based measures [Pope et al., 1998; Riley, 2000; Kautter and Pope, 2004]. In addition, performing instrumental activities of daily living (IADLs, i.e., using the phone, doing light housework, doing heavy housework, making meals, shopping and managing money) is more cognitively demanding than performing ADLs, and is directly impacted by the progression of AD, making IADLs less suitable for assessing comorbidities. Therefore, the frailty adjuster in our study uses a scale based on the count of ADL impairments, and was categorized as none, 1-2 (low), 3-4 (moderate), and 5-6 (high). For each activity, individuals were asked whether they had any difficulty performing the activity, received help with the activity, needed supervision with the activity, or were unable to perform the activity because of health problems. An individual was coded as impaired for an ADL if he or she needed any assistance doing (i.e., received help or needed supervision) or could not perform the activity.

Prior expenditures

Expenditures incurred in year *t* are highly correlated with expenditures in year *t+1* [Garber et al., 1997; Pope et al., 1998]. The distributions of overall and drug expenditures were right-skewed in our sample: less than 1% of the sample had no overall expenditures and less than 4% had no drug expenditures. We modeled expenditures in year *t* as continuous variables on their original scale, and used the generalized linear model (GLM,

described in the next section) as an alternative to transforming the non-normally distributed data.

Combined models

We also compared the performance of combined models, which incorporate the frailty adjuster or/and prior expenditures in the diagnosis-based measures (i.e., CMS-HCC, CDPSM and CCI), against the performance of single-measure models.

Analysis

A Pearson correlation coefficient matrix was used to evaluate the correlation among various risk adjustment measures. Ordinary least squares (OLS) regressions were employed to predict expenditures in year *t+1* using each risk adjustment measure plus age-gender categories in year *t*. Because the expenditure data were non-normally distributed, we also performed the generalized linear model (GLM) with a gamma variance and a log link function, which models expenditures in their natural scale rather than in a transformed scale (e.g., log transformation) [Shwartz and Ash, 2003]. The full models are described in Equations 6.1 and 6.2.

OLS: $E[Expenditure_{t+1}] = \beta_1 Age_{it} + \beta_2 Sex_{it} + \beta_3 Diagnosis_{it} + \beta_4 ADL_{it} + \beta_5 Expenditure_{it}$ (Equation 6.1)

GLM: $g[E(Expenditure_{t+1})] = \beta_1 Age_{it} + \beta_2 Sex_{it} + \beta_3 Diagnosis_{it} + \beta_4 ADL_{it} + \beta_5 Expenditure_{it}$ (Equation 6.2)

Since this study focuses on an AD-specific population with a sample size of 671 unique individuals, we were unable to use a split-sample design to evaluate predictive accuracy, in which a small randomly selected sample of the total study population is withheld for model validation [Ellis et al., 1996; Fishman et al., 2003; Meenan et al., 2003; Pacala et al., 2003; Powers et al., 2005]. Therefore, alternative statistical evaluation criteria were adopted. We used adjusted R^2 from OLS models to demonstrate the proportion of total variance in the dependent variable (e.g., expenditures) accounted for by the risk adjustment model [Ellis et

al., 1996]. For GLMs, we compared the log likelihood values across models, with higher numbers indicating better model fit.

Although adjusted R^2 and log likelihood values provide a summary measure of overall prediction, these statistics give little information about how well a model discriminates between high- and low-expenditure cases [Shwartz and Ash, 2003]. Therefore, an individual's actual expenditures in the prediction year (i.e., year *t+1*) were categorized into quartiles; then, predictive ratios (i.e., predicted expenditures divided by actual expenditures) were calculated within each quartile [Ash et al., 2000; Cucciare and O'Donohue, 2006]. If the model performs well for a population, its predictive ratio is close to one. In a prospective payment system, a predictive ratio close to one indicates that aggregate payments under the risk adjustment model are equivalent to payments under the fee-for-service [Ellis et al., 1996]. A ratio greater than one indicates that payments are set higher than the actual expenditures incurred, whereas a ratio less than one represents under-payment for the actual expenditures.

Predictive model performance was examined further using receiver operating characteristic (ROC) curves, representing how well the model classifies individuals by illustrating the tradeoff between true positive (sensitivity) and false positive (1-specificity) for a certain high-expenditure threshold [Meenan et al., 1999]. This is done by rank-ordering actual and predicted expenditures from high to low, and setting a pre-determined percentage threshold (e.g., top 10%) within each expenditure distribution to define "true" high-expenditure cases [Meenan et al., 1999; Meenan et al., 2003; Weiner, 2003]. A *c*-statistic representing discrimination power then can be calculated using a nonparametric trapezoidal method to approximate the area under the ROC curve [Shwartz and Ash, 2003]. A value of 0.5 indicates no ability to discriminate; higher values between 0.5-1.0 indicate a better fit.

6.6 Results

The average age of the sample was 81.2 years, and 60% were female (Table 6.1). All three diagnosis-based risk adjustment measures indicated a substantial burden of comorbidities: the average CMS-HCC, CDPSM, and CCI scores were 1.6, 1.6, and 2.8, respectively. The burden of comorbidities among our AD sample was higher compared with a sample of Veterans regularly using primary care [Maciejewski et al., 2005]: the average DCG, CDPS, and CCI scores were 1.3, 1.9, and 0.2, respectively. On average, individuals with AD had 1.4 ADL limitations: approximately 60% reported no ADL impairment, whereas 14% reported impairment for at least five of six ADLs. Average total expenditures in prediction year *t+1* were $18,736 (standard deviation [s.d.]= $31,163); 11 of the 671 individuals (1.6%) had total expenditures exceeding $100,000. Average prescription drug expenditures were $2,562 (s.d.= $2,511) in the prediction year; 15 (2.2%) spent more than $10,000 on prescription medicine. The average per capita overall and drug expenditures among seniors with AD were substantially higher their counterparts without AD ($10,371 for total and $1,681 for drug expenditures, results not shown).

Table 6.1: Descriptive statistics of sample of individuals with Alzheimer's disease

Characteristics	Mean (s.d.)	Minimum	Maximum
Age, years	81.2 (7.1)	65	101
65-69	6.7%	--	--
70-74	13.1%	--	--
75-79	18.0%	--	--
80-84	29.4%	--	--
85+	32.8%	--	--
Male	40.1%	--	--
CCI	2.8 (2.6)	0	15.0
CMS-HCC	1.6 (1.2)	0.3	8.3
CDPSM	1.6 (1.2)	0.2	7.5
Count of ADL limitations	1.4 (2.0)	0	6.0
0	59.5%	--	--
1-2	16.0%	--	--
3-4	11.0%	--	--
5-6	13.6%	--	--
Total expenditures, 2007$			
Year 1	$16,899 ($21,282)	0	$219,992
Year 2	$18,736 ($31,163)	0	$364,116
Drug expenditures, 2007$			
Year 1	$2,401 ($2,263)	0	$24,832
Year 2	$2,562 ($2,511)	0	$27,553

ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category; IADL: independent activity of daily living; s.d.: standard deviation

The correlation matrix (Table 6.2) revealed that the three diagnosis-based risk adjustment measures had moderate-to-high correlations, especially the CMS-HCC and the CDPSM (r=0.93). This high correlation is likely due to similar approaches used in both measures to define diagnoses, to assign disease diagnoses, and to aggregate disease categories into larger areas according to body system or type of disease [Kronick et al., 2002]. The CMS-HCC and the CDPSM also were strongly correlated with total expenditures in year *t* (r=0.67 and 0.68, respectively). The frailty adjuster measured by ADL limitations was only moderately correlated with the CMS-HCC and total expenditures in the base year, but not with other measures. Most other correlations were low, ranging between 0.1 and 0.3.

Table 6.2: Correlation matrix of risk adjustment measures

	CCI	CMS-HCC	CDPSM	ADL limitations	Prior total expenditures
CMS-HCC	0.29**				
CDPSM	0.31**	0.93**			
ADL limitations	0.03	0.29**	0.26**		
Prior total expenditures	0.21**	0.67**	0.68**	0.29**	
Prior prescription drug expenditures	0.12**	0.16**	0.17**	0.09*	0.23**

p<0.05, ** p< 0.01
ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category

Adjusted R^2 values for prospective OLS models and log likelihood values of the GLMs are reported in Table 6.3. Diagnosis-based risk adjustment models explained more variance in total expenditures than in drug expenditures. Age and gender alone explained less than 1% of the total expenditure variation. Adding diagnosis-based measures to the demographic model substantially increased the explanatory power. In predicting total expenditures, the CCI had a higher adjusted R^2 (17.0%) compared to the CDPSM (15.5%) and the CMS-HCC (12.1%). Prior expenditures were more predictive (adjusted R^2=20.4%) than any other single measures. Incorporating the frailty adjuster modestly increased the adjusted R^2 of the diagnosis-based models, with the CCI remaining the most predictive (18.5%), followed by the CDPSM (15.6%) and CMS-HCC (14.2%), but not the prior-expenditure model. Adding prior expenditures to diagnosis-based measures with frailty adjustment increased the adjusted R^2 to 21.1%-31.1%. The log likelihood values in GLMs revealed similar results as with adjusted R^2 statistics, although the CCI had slightly higher log likelihood values than prior expenditures, regardless of whether frailty adjustment was used.

Diagnoses were not predictive of future prescription drug expenditures (adjusted R^2=1.4%-2.3%). Incorporating the frailty adjuster even decreased the explanatory power. Adding prior drug expenditures increased the adjusted R^2 to approximately 37%. The log

likelihood values in GLMs also indicated that models incorporating prior drug expenditures had a better fit.

Table 6.3: Adjusted R^2 and log likelihood values of risk adjustment measures in prospective expenditure models

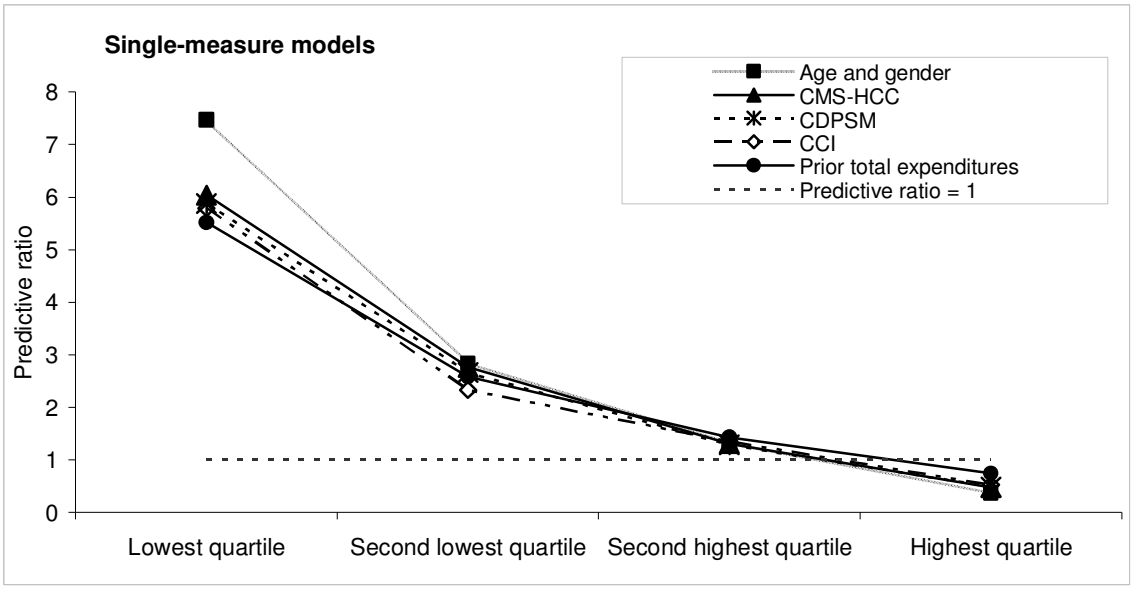
	Adjusted R-squared[1] x 100%		Log likelihood value[2]	
Risk adjustment measures	Total expenditures	Prescription drug expenditures	Total expenditures	Prescription drug expenditures
Age and gender	0.02	0.9	-7203.3	-5724.7
Diagnosis-based measures				
CMS-HCC	12.1	1.4	-7130.8	-5722.6
CDPSM	15.5	1.7	-7127.2	-5721.3
CCI	17.0	2.3	-7110.3	-5718.3
Frailty adjuster				
Counts of ADL limitations[3]	2.0	0.5	-7185.1	-5724.5
Prior expenditures[4]	20.4	37.1	-7119.4	-5599.2
Combined models				
CMS-HCC + count of ADL limitations	14.2	1.0	-7124.6	-5722.6
CDPSM + count of ADL limitations	15.6	1.3	-7119.5	-5721.3
CCI + count of ADL limitations	18.5	1.9	-7093.7	-5718.2
Prior expenditures + count of ADL limitations	20.3	37.3	-7116.5	-5597.0
CMS-HCC + count of ADL limitations + prior expenditures	21.1	37.2	-7108.9	-5597.0
CDPSM + count of ADL limitations + prior expenditures	21.6	37.2	-7105.3	-5596.8
CCI + count of ADL limitations + prior expenditures	31.1	37.5	-7054.5	-5596.1

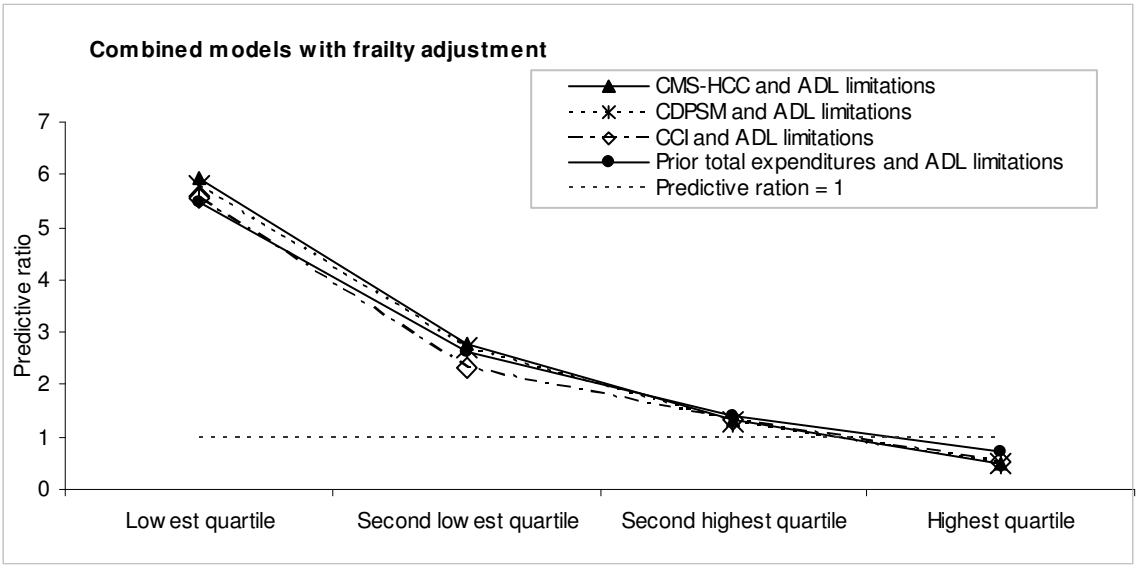
[1] Adjusted R-squared is obtained from ordinary least square regression.
[2] Log likelihood is obtained from generalized linear model with a gamma variance and a log link function.
[3] Counts of ADL limitations were categorized as none, 1-2, 3-4, and 5-6.
[4] Prior total expenditures were used to predict subsequent-year total expenditures; prior drug expenditures were used to predict subsequent-year drug expenditures.
ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category.

Figures 6.2 and 6.3 illustrate predictive ratios (i.e., predicted-to-actual expenditures) by expenditure quartile of overall and drug expenditures. For total expenditures, all single-measure and combined models substantially under-predicted expenditures in the highest quartile, and over-predicted those in the lowest quartile. Compared to the CMS-HCC and the CDPSM, the CCI exhibited less deviation from the reference line indicating risk-adjusted capitation payments equivalent to payments under the fee-for-service (i.e., predictive ratio=1) in the lowest, second lowest and highest quartiles. All three diagnosis-based measures under-predicted actual total expenditures in the highest quartile

by approximately 50%, and over-predicted those in the lowest quartile by six fold. Inclusion of frailty adjustment improved overall explanatory power but only modestly improved the predictive accuracy in each expenditure quartile. With frailty adjustment, the CCI also exhibited less deviation from the reference line compared with the CMS-HCC and the CDPSM. Prior expenditures were more accurate in predicting individuals in the lowest and in the highest quartiles compared with diagnosis-based measures, regardless of whether frailty adjustment was applied.

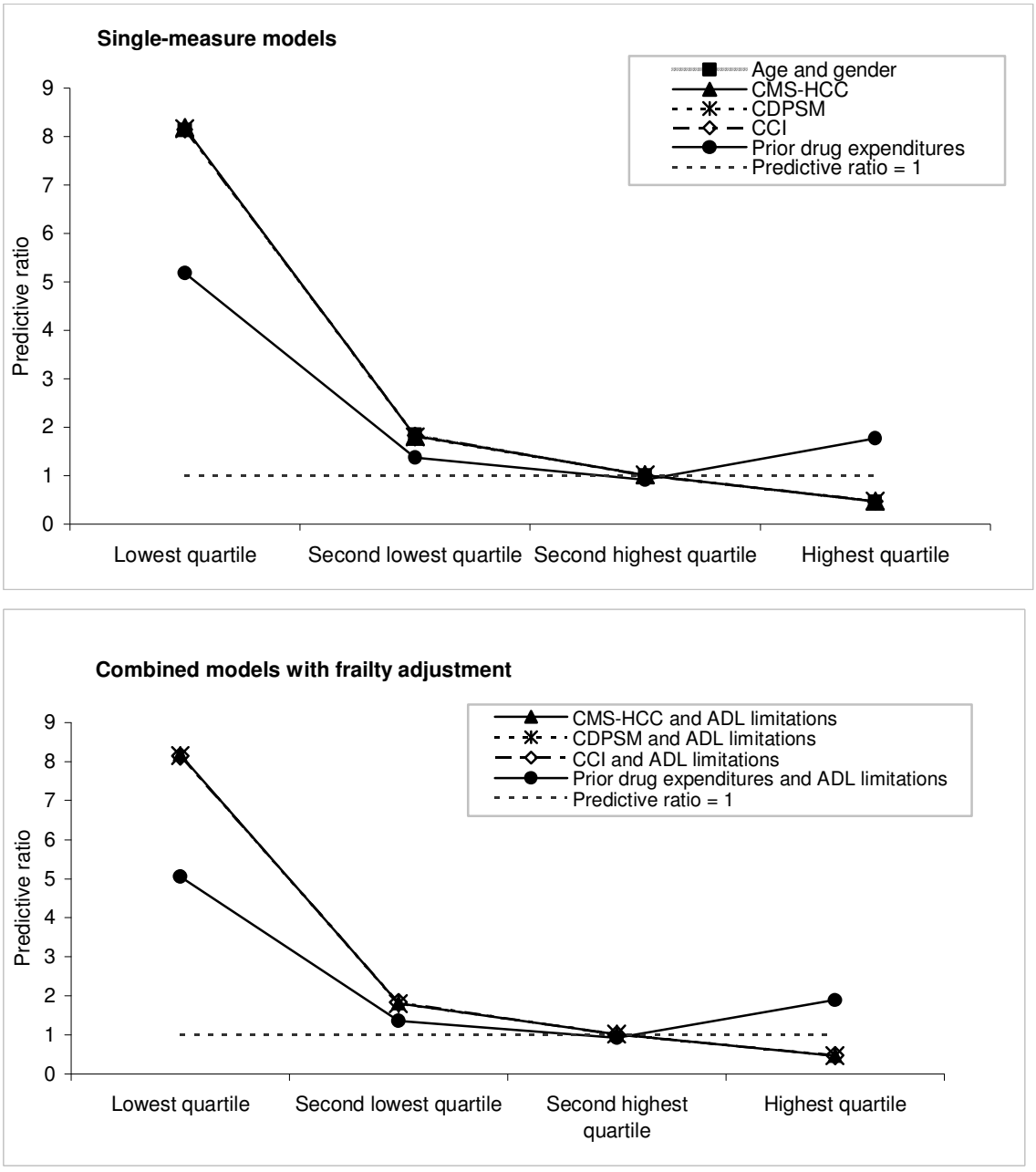
For prescription drug expenditures, all three diagnosis-based measures performed similarly: they under-predicted actual drug expenditures among individuals in the highest quartile by approximately 50%, and over-predicted those in the lowest quartile by eight fold, a greater over-prediction than for total expenditures. Prior drug expenditures were more accurate than any other measures but tended to over-predict across quartiles. Inclusion of ADL limitations did not improve the predictive accuracy in each drug expenditure quartile.





ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category.

Figure 6.2: Predictive ratios of risk adjustment measures in predicting total expenditures



ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category.

Figure 6.3: Predictive ratios of risk adjustment measures in predicting prescription drug expenditures

A model's ability to predict being in the top 10% of the expenditure distribution (i.e., "high-expenditure outliers") was assessed by the area under the ROC curves (Table 6.4). In predicting total expenditures, demographic characteristic had no ability to discriminate

(*c*=0.509). Adding diagnoses raised the *c*-statistics to 0.667 (CCI), 0.645 (CDPSM), and 0.628 (CMS-HCC). Inclusion of the frailty adjuster slightly improved the discrimination power of the CDPSM (*c*=0.650) and the CMS-HCC (*c*=0.635), but not the CCI (*c*=0.664). Prior expenditures were more predictive (*c*=0.703) than any other single-measure models or combined models with frailty adjustment.

For prescription drug expenditures, diagnoses had only modest discrimination power of classifying individuals into high-expenditure category. Adding frailty adjustment slightly increased the *c*-statistics of the CCI (0.618) and the CDPSM (0.564), but not the CMS-HCC (0.561). Prior drug expenditures had good discrimination power (*c*=0.848) and outperformed any other single-measure models or combined models with frailty adjustment.

Table 6.4: Area under the receiver operating characteristic curves in prospective expenditure models

Risk adjustment measures	10% high-expenditure threshold	
	Total expenditures	Drug expenditures
Age and gender	0.509	0.587
Diagnosis-based measures		
CMS-HCC	0.628	0.600
CDPSM	0.645	0.561
CCI	0.667	0.614
Frailty adjuster		
Counts of ADL limitations[1]	0.521	0.585
Prior expenditures	0.703	0.848
Combined models		
CMS-HCC + count of ADL limitations	0.635	0.561
CDPSM + count of ADL limitations	0.650	0.564
CCI + count of ADL limitations	0.664	0.618
Prior expenditures + count of ADL limitations	0.685	0.847
CMS-HCC + count of ADL limitations + prior expenditures	0.686	0.846
CDPSM + count of ADL limitations + prior expenditures	0.693	0.845
CCI + count of ADL limitations + prior expenditures	0.713	0.848

[1] Counts of ADL limitations were categorized as none, 1-2, 3-4, and 5-6.
ADL: activity of daily living; CCI: Charlson Comorbidity Index; CDPSM: Chronic Illness and Disability Payment System-Medicare; CMS-HCC: Centers for Medicare and Medicaid version of the Diagnostic Cost Group-Hierarchical Condition Category.

6.6 Discussion

This study compared the performance of five different risk adjustment measures in predicting overall and prescription drug expenditures for Medicare beneficiaries with Alzheimer's disease. Our analysis showed that, in single-measure, diagnosis-based models, the CCI outperformed the CMS-HCC and the CDPSM in predicting both total and drug expenditures. CCI remained the best-performing measure when adding frailty adjustment based on counts of ADL limitations. The frailty adjuster increased overall explanatory power and modestly improved the predictive accuracy in each expenditure quartile for predicting total expenditures but not for drug expenditures. In high-expenditure outlier analysis, the CCI exhibited greater discrimination power than the CMS-HCC and the CDPSM, regardless of whether frailty adjustment was applied.

Our data indicated that, models incorporating prior expenditures generally had greater prediction power and predictive accuracy, especially in predicting drug expenditures. Past use explained more than 37% of the total variations in drug expenditures. Only prior expenditures, but not diagnoses, predicted drug expenditures with any accuracy, and can discriminate between high- and low-expenditure individuals. The difference in accuracy between prior expenditures and diagnoses was greatest in individuals with the highest drug expenditures (i.e., the top 25%). This phenomenon is likely due to a strong degree of persistence in prescription drug expenditures among the elderly, in which high-expenditure users remain in the upper percentiles and low-expenditure individuals remain in the bottom of the expenditure distribution in the next year [Stuart et al., 1991; Coulson and Stuart, 1992; Wrobel et al., 2003; Lin, 2008b]. However, inclusion of prior expenditures in a prospective risk adjustment model to set payment might not be practical because of the perverse incentives created [Ash et al., 2000; Greenwald, 2000; Ash et al., 2001]. For instance, models that pay more for heavy users may encourage both appropriate and inappropriate or unnecessary services [Ash et al., 2001].

On the other hand, combining expenditure and diagnosis data may be more powerful and more operationally useful for identifying beneficiaries who may be at risk for higher future expenditures. Particularly for AD, MCOs have many opportunities to mitigate costs or change care patterns through disease management, including use of appropriate medications (e.g., cholinesterase inhibitors), discharge planning, education, counseling, and support for caregivers [Fillit et al., 2002b; Kaufer et al., 2005], which may delay time to nursing home care. Better care coordination, such as medication management by nurses or pharmacists, is needed for improving adherence to drug therapies for concomitant conditions in addition to AD.

Without risk adjustment, payers using capitation payment methods are likely to overpay (or underpay) providers for healthier plan enrollees (or sicker groups). If payers do not reimburse more money to providers who serve enrollees with above-average levels of health care needs, plans will be penalized for enrolling sicker patients and quality of care may be jeopardized [Kronick et al., 2000]. Diagnosis-based measures, such as the PIPDCG and the DCG-HCC, have been shown to under-predict (or over-predict) average expenditures for Medicare beneficiaries with (or without) ADL impairments [Riley, 2000]. The degree of under-prediction has been shown to increase with the number of ADL limitations, which are associated with Medicare expenditures, but not fully captured in demographic characteristics or diagnosis profiles. Beginning in 2004, frailty-adjusted Medicare payments were applied to some MCOs, such as Program of All Inclusive Care for the Elderly (PACE), Wisconsin Partnership Program, and Minnesota Senior Health Options, that specialize in providing care to the community-residing frail elderly [Kautter and Pope, 2004]. For individuals with AD, progressive functional decline and thus frailty is well documented. Using data from the 1994 National Long Term Care Survey, Taylor and colleagues found that functional impairment was more predictive than cognitive status of total expenditures for individuals with AD and related dementias, and that risk adjustment measures should account for functional disability

and comorbidity rather than just AD status [Taylor et al., 2001]. However, under the current payment system, individuals with AD will not qualify for an enhanced capitation payment rate if they are enrolled in traditional Medicare MCOs. Our data showed that with frailty adjustment based on counts of ADL limitations, the best-performing model, CCI, had an adjusted R^2 of 18.5%. Incorporating ADL limitations increased the explanatory power of the CMS-HCC model by 17%, an amount greater than the 11% improvement found by including this adjustment in developing payment for the general Medicare beneficiaries [Kautter and Pope, 2004]. Our results support including ADL limitations as a frailty adjuster in a diagnosis-based risk adjustment model until better measures of risk can be developed. Nevertheless, the acquisition, cleaning, and use of administrative data are often time-consuming and costly [Powers et al., 2005]. Unlike in long-term-care settings in which functional ability is collected routinely as part of the Minimum Data Set Nursing Home Assessment Record [CMS, 2008], payers should assess the time and cost of collecting additional information on functional ability for community-dwelling individuals, as this information is not readily available in administrative claims data.

Risk adjustment models differ from each other in the characteristics used to explain expenditures and how this information is organized in the classification system [Greenwald, 2000]. Compared with the other two diagnosis-based measures, the CCI consists of fewer disease categories and a simpler scoring algorithm. The CCI has been validated in many studies to predict mortality outcome [Schneeweiss and Maclure, 2000; Needham et al., 2005], but this methodology has not been thoroughly explored to predict expenditures. Surprisingly, the CCI outperformed the CMS-HCC and the CDPSM in many instances in our AD sample. In predicting total expenditures, the CCI produced an adjusted R^2 of 17%, comparable to the 20% threshold found in most prospective models based on diagnosis information or on pharmacy claims [Newhouse et al., 1989; Shwartz and Ash, 2003]. With frailty adjustment, the *c*-statistics of diagnosis-based models ranged between 0.64 and 0.66 in our AD sample.

The results are comparable to an assessment of risk adjustment models based on diagnosis, pharmacy claims, or prior expenditures among elderly managed care beneficiaries in which the *c*-statistics ranged between 0.65 and 0.69 in models given a highly stringent 1% high-expenditure threshold [Meenan et al., 2003]. We also validated the results by excluding individuals with extremely high overall expenditures (>$100,000) and extremely drug expenditures (>$10,000), and found little effect on our conclusions. Our findings support that, properly combining frailty adjustment and comorbidity profiles, risk adjustment measures appear promising in predicting health expenditures even in a population with substantial disabilities.

In interpreting the results presented here, several notes of caution are important. First, the study sample excluded individuals residing in skilled nursing facilities who generally have more functional impairments and more comorbidities than community-dwelling individuals with AD [Hill et al., 2006]. In the future, diagnosis-based models with frailty adjustment will require further review and refinement for risk adjustment to be applicable to long-term-care settings. For example, an admission profile, such as the pressure ulcer prevalence rate, may be adjusted to recognize that some facilities admit sicker individuals [Berlowitz and Rosen, 2003]. The study sample also excluded individuals with any managed care participation because their medical claims and diagnoses were not available in the MCBS. In general, these individuals were somewhat younger and healthier and had lower average overall and drug expenditures than our study sample. Future investigation should evaluate the performance of risk adjustment models for certain MCO subpopulations, especially individuals with functional impairments.

According to the Medicare Part D risk adjustment model (i.e., RxHCC) [CMS, 2006a], dementia with depression or behavioral disturbance is associated with $1,104 (2006 dollars) increase in expenditures for community residents. However, we did not have ambulatory pharmacy claims data to determine whether pharmacy-based models, such as the Chronic

Disease Score [Von Korff et al., 1992] and the RxRisk model [Fishman et al., 2003], would have improved predictive accuracy, especially in predicting drug expenditures. Future research is needed to evaluate the performance of pharmacy-based measures in order to improve Part D payment methodology. For seniors with AD, assessment of pharmacy-based risk adjustment models can help policymakers to design formularies that offer better and more comprehensive drug coverage to meet their different needs, such as medications for treating common comorbidities including psychiatric disorders and diabetes [Rice et al., 2001].

Another issue concerns the quality of self- or proxy-reported functional status. In our sample, nearly 50% of proxy response occurred on behalf of individuals with AD. Although self-report may have greater errors in individuals with cognitive impairment relative to the general population, it is not clear whether proxy reports are systematically more or less accurate than self reports. Individuals with dementia may overestimate their cognitive and functional abilities compared with proxy ratings [Koss et al., 1993; Tierney et al., 1996; Farias et al., 2005]. On the other hand, greater caregiver burden may lead to overstated functional impairments of individuals with cognitive impairment [Rothman et al., 1991; Long et al., 1998]. In the future it will be important to investigate the validity of proxy reports in order to apply accurately the frailty adjuster based on functional ability.

6.6 Conclusion

In single-measure models, CCI outperformed CMS-HCC and CDPSM in predicting both total and drug expenditures among our sample of community-dwelling Medicare beneficiaries with AD. A frailty adjuster based on count of ADL limitations improved overall prediction and predictive accuracy of the diagnosis-based models, especially the CMS-HCC specification currently used by Medicare. No model except past use predicted drug

expenditures well. Future research is needed to evaluate the performance of risk adjustment measures based on ambulatory pharmacy data.

CHAPTER 7

STUDY LIMITATIONS AND FUTURE RESEARCH AGENDA

7.1 Summary of Findings

This dissertation examined the dynamics of overall health expenditures and prescription drug expenditures in Medicare beneficiaries with Alzheimer's disease (AD), a population with progressive disabilities and associated with high expenditures. Previous cost-of-illness studies on AD yield great variation in cost and prevalence estimates. As well, little is known about the characteristics of individuals with AD who accrue high and persistently high expenditures. This dissertation sought to characterize their expenditure profiles prospectively using data from the 1999-2004 Medicare Current Beneficiary Survey (MCBS) linked with Medicare claims files. In addition, this study was the first, to our knowledge, to compare the performance of various risk adjustment measures to predict overall and prescription drug expenditures in the AD population. The findings of this dissertation may help payers to compensate health plans fairly for the expected expenditures associated with the disease burden of their enrollees with AD.

The case-finding analysis presented in Chapter 4 identified the most sensitive and specific alternatives for identifying who has AD. The use of survey report, diagnosis listed in medical claims, use of Alzheimer-specific prescription medicine, or some combination of these three definitions identified different subsets of individuals with AD. Using diagnosis as a "gold-standard", as is typical for other diseases and conditions, survey report was able to identify correctly individuals who do not have AD (i.e., highly specific) without losing the ability to identify correctly individuals who have AD (i.e., sensitivity) found when Alzheimer's

medication use is used as a marker for the disease. As a consequence of the different disease definitions employed, the expenditure estimates for individuals with AD varied widely. 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. Thus, one should exercise caution when interpreting current cost-of-illness studies and in applying these estimates to policy initiatives. In observational studies, the development of a comprehensive case definition, which incorporates all available markers for AD to the researcher, is a crucial first step to assessing the health care needs for persons with AD.

Against the background of relatively high expenditures by Medicare beneficiaries with AD shown in Chapter 4, Chapter 5 provided information on the dynamic distribution of overall and prescription drug expenditures in this population. Specifically, individuals in the top 10% of the expenditure distribution (i.e., those with the highest expenditures) accounted for 38%-47% of overall health expenditures and 31%-36% of overall drug expenditures, depending on the study year. Thus, health expenditures (both for drugs and other types of care) among Medicare beneficiaries with AD were highly concentrated in a relatively small proportion of individuals. A quarter of the highest-spending 10% for total health expenditures remained in the top decile in the next year, whereas 21% of them moved to the bottom half of the distribution in the subsequent year. Half of the highest 10% with drug expenditures retained this ranking but only 9% became the bottom 50% in a second year, indicating that prescription drug expenditures were more persistent than overall expenditures. Prior expenditures and comorbidities, but not functional status, were strong predictors of the level of future expenditures. The potential for high expenditures by a relatively small group of individuals in subsequent years, coupled with the likelihood for growth of AD among the aging population, suggests the need for providing evidence that additional care coordination and disease management programs may be necessary to improve health outcomes and to control costs.

Building upon the findings in Chapter 5 that expenditures among individuals with AD were correlated with personal characteristics and prior expenditures and, therefore, can be predicted, Chapter 6 compared the performance of various prospective risk adjustment models of predicting overall and prescription drug expenditures. The specific risk adjustment measures evaluated for this sample were constructed from diagnoses in medical claims (i.e., Charlson Comorbidity Index [CCI], Hierarchical Condition Category model used by Medicare [CMS-HCC], and Chronic Illness and Disability Payment System-Medicare version [CDPSM]), functional status (i.e., limitations of activities of daily living [ADLs]) as reported in the MCBS, prior expenditures, and some combination thereof. In line with the evidence found in the general Medicare population, this analysis demonstrated that total expenditure models with multiple risk adjustment measures outperformed single-measure models in this AD sample. The diagnosis-based models under-predicted the actual total expenditures among individuals in the highest quartile by approximately 50%, and over-predicted those in the lowest quartile by 6-fold. The frailty adjuster based on functional ability had very limited predictive power itself, whereas it improved overall prediction and predictive accuracy on the diagnosis-based models. Inclusion of the frailty adjuster in the diagnosis-based models in predicting total expenditures raised the adjusted R^2 to 19%, very similar to the 20% threshold found in most prospective models based on diagnosis information or on pharmacy claims [Newhouse et al., 1989; Shwartz and Ash, 2003]. With frailty adjustment, the *c*-statistics of diagnosis-based models ranged between 0.63 and 0.67 in our AD sample. The results are comparable to an assessment of risk adjustment models based on diagnosis, pharmacy claims, or prior expenditures among elderly managed care beneficiaries in which the *c*-statistics ranged between 0.65 and 0.69 in models given a 1% high-expenditure threshold [Meenan et al., 2003]. However, for drug expenditures, only prior expenditures, but not comorbidities or functional status, predicted well. Incorporating prior drug expenditures increased the adjusted

R^2 to 37%, which is still lower than the explanatory power observed in a nationally representative sample of Medicare beneficiaries (adjusted R^2=55%) [Wrobel et al., 2003].

Based on these results, expenditure analysis for individuals with AD using observational data should incorporate all sources of definitions available to researchers as a first step to case ascertainment. If only diagnosis code data are available, as in administrative claim data, researchers should exercise substantial caution when reporting prevalence and expenditure data given the proportion of missed AD cases we noted in this study. The development of a comprehensive case definition that is reasonably sensitive yet specific is important for identifying individuals at different phases of AD and to assessing their health care needs. For persons with a chronic condition like AD that places substantial burdens on both the individual and the health care system, there is a critical need to understand the dynamics of their expenditure profiles. This dissertation found that overall health care and drug expenditures were highly concentrated and persistent over a two-year period in this AD population. Prior expenditures and comorbidities, were highly predictive of an individual's percentile position in the subsequent-year expenditure distribution, much stronger than the effects of demographic characteristics and functional ability. Prospective risk adjustment models combining multiple measures, such as functional ability and comorbidity profiles, improve predictive accuracy of current diagnosis-based capitation payment methods. Therefore, with proper organization of information, risk adjustment measures appear promising in predicting health expenditures even in a population with substantial disabilities, such as individuals with AD. Predictive models identifying seniors with AD who are at high risk of accruing high expenditures can help MCOs to mitigate costs or change care patterns through disease management, including use of appropriate medications (e.g., cholinesterase inhibitors), discharge planning, education, counseling, and support for caregivers [Fillit et al., 2002b; Kaufer et al., 2005], which may delay time to nursing home care. Better care

coordination, such as medication management by nurses or pharmacists, could improve adherence to drug therapies for concomitant conditions in addition to AD.

7.2 Study Limitations

Although these analyses provide important contributions to the study of dynamics of overall and drug expenditures among individuals with AD, there are limitations that present an opportunity for future investigation. First, the findings presented in Chapter 4 represent the necessary first step, not a definitive or absolute analysis, towards appropriate case ascertainment for identifying all individuals with AD. In our analysis, we were unable to rule out potentially false-positive cases, such as individuals who report being told by their physicians that they have AD but who do not have an AD diagnosis in their medical claims and/or who are not taking Alzheimer-specific medications (i.e., acetylcholinesterase inhibitors or memantine). They may be individuals with mild cognitive impairment (MCI) or early-stage AD who are under-diagnosed or uncoded in claims data because of social resistance to ageism and disabling stigma [Fillit, 2000; Rice et al., 2001]. It is also possible that AD is under-coded because the reimbursement scheme provides little financial incentive for coding AD as the primary diagnosis and instead encourages coding comorbidities, such as aspiration pneumonia as the primary diagnosis, to enhance reimbursement [Newcomer et al., 1999; Fillit, 2000; Brummel-Smith, 2001; Rice et al., 2001]. Moreover, an estimated 50% of individuals with AD are diagnosed, and only half of those diagnosed are being treated [Evans, 1990]. Individuals with AD may not take acetylcholinesterase inhibitors or memantine because these drugs may temporarily delay memory decline but none of them is known to stop the underlying degeneration of brain cells [Alzheimer's Association, 2007]. The lack of a uniformly accepted definition or a gold-standard diagnostic test result poses a challenge to verifying the validity of survey report, diagnosis, and medication use that we employed to define AD in the MCBS.

Second, although our sample was drawn from a nationally representative Medicare sample, use of secondary data brings limitations. There might be errors in data collection, editing and imputation, which are difficult to evaluate but can result in biased estimates. Particularly, the reliability of self report by persons with AD and proxy response may be imperfect. Our data revealed that proxy use was much higher in individuals with AD than in those without, probably due to great disabilities among this population. A number of studies have shown that, among individuals with dementia, self report often differs substantially from information collected from caregivers [DeBettignies et al., 1990; Farias et al., 2005], whereas it is unclear whether proxy reports are systematically more or less accurate than self reports. For instance, individuals with AD may experience a great loss of insight and thus overestimate their cognitive and functional abilities compared with proxy ratings [Koss et al., 1993; Tierney et al., 1996; Farias et al., 2005]. On the other hand, greater caregiver burden may lead to overstated functional impairments of individuals with cognitive impairment [Rothman et al., 1991; Long et al., 1998]. Therefore, functional status collected from self report or proxy response should be validated, especially in a subpopulation with cognitive and memory loss, to enhance the performance of frailty adjustment on risk-adjusted capitation payments.

Third, we had to limit our sample to individuals with complete claims records in the MCBS and excluded those enrolled in managed care because these plans do not submit detailed claims to the Centers for Medicare and Medicaid (CMS), and thus we were unable to construct their comorbid condition profiles using diagnoses listed in the medical claims. We examined these individuals separately using survey data and found that they were somewhat younger and healthier and had lower average overall and drug expenditures than our study sample. Therefore, one should be cautious in interpreting the results as these findings may not be generalizable to managed care enrollees.

Fourth, our sample also excluded residents in long-term-care facilities defined as individuals who were institutionalized for more than 90 days at a time during a year [CMS, 2006c]. Evidence has shown that individuals with AD residing in skilled nursing facilities tend to have more functional impairments and more comorbidities than community residents [Hill et al., 2006]. They may be individuals with late-stage AD who require a significant amount of caregiving and supervision each day, rather than the ones who are well enough to live at home. Therefore, our findings may not be generalizable to individuals with AD living in long-term-care facilities.

On the other hand, the act of institutionalization is a choice made by individuals/caregivers but may not necessarily define whether she or he *requires* such intensive care [Caro et al., 2001]. Therefore, our community-dwelling sample may be a heterogeneous group with individuals at different phases of AD and with various health care needs. Because the MCBS does not provide objective measures of severity of AD, such as the modified Mini-Mental State Examination [Folstein et al., 1975], our study was unable to examine the relationship between severity and expenditures. Nor were we able to determine whether individuals incurred persistently high expenditures was due to more advanced AD or more severe comorbid conditions, or the synergetic effect of both.

Moreover, criticism also can be levied about the lack of automated ambulatory pharmacy data in our study, although drug expenditures were available in the MCBS based on survey information to impute payments from all sources, including Medicare, Medicaid, Medicare MCOs, private MCOs, the Veterans Administration, employer-sponsored private insurance, individually-purchased private insurance, private insurance from an unknown source, out-of-pocket payments, and public health plans other than Medicare or Medicaid [CMS, 2003]. In the case-finding analysis, information on prescription drug use was ascertained from survey data rather than actual claims records. However, prescriptions filled several months earlier may not be a salient event in the typical respondent's memory

[England et al., 1994], especially those with memory impairments like persons with AD. Individuals may have difficulty saving all prescription containers over the typical four-month span between interviews. As a result, the prevalence of AD based on prescription medication use was likely underestimated due to under-reporting of medication in addition to any under-treatment that may occur. Therefore, pharmacy data should be used only as a supplement definition for case ascertainment rather than a definitive measure for AD.

Lastly, in the risk adjustment analysis, no model except prior expenditures predicted drug expenditures well. Nevertheless, inclusion of past use in a prospective risk adjustment model might not be practical because of the perverse incentives created for payment purposes [Ash et al., 2000; Greenwald, 2000; Ash et al., 2001]. For instance, models that pay more for expensive, heavy users may encourage both appropriate and inappropriate or unnecessary services [Ash et al., 2001]. Nevertheless, during the time frame of our study, Medicare Part D benefit was not yet available to its beneficiaries, and the MCBS does not supply any drug claims for us to construct pharmacy-based models, such as the Chronic Disease Score [Von Korff et al., 1992] and RxRisk [Fishman et al., 2003]. As a result, we were unable to determine whether pharmacy-based measures would have improved predictive accuracy on total and drug expenditure models.

7.3 Future Research Agenda

The conclusions as well as the limitations of this study suggest possible avenues for future research, including the use of alternative data sources and expansion of study population. Specific directions for future investigation are outlined in Table 7.1.

Table 7.1: Directions for future research

Case definition	▪ Validate self- or proxy-reported AD and AD diagnosis listed in medical claims with objective measures
Utilization patterns	▪ Examine the type of service use and treatment patterns between individuals with persistently high and low expenditures ▪ Evaluate managed care enrollees' expenditure profiles and utilization patterns
Risk adjustment measures	▪ Validate information on functional status as a frailty adjuster ▪ Evaluate the performance of pharmacy-based measures
Long-term-care population	▪ Estimate expenditure concentration and persistence in institutionalized individuals with AD ▪ Study risk adjustment models in long-term-care settings

One of the most important agenda items lies in continuing the development of a sensitive and specific definition for AD in observational studies. Future research should seek to validate self- or proxy-reported AD and AD diagnosis listed in medical claims with objective measures, such as cognitive function assessment, neurological examinations, apolipoprotein E genotype testing (i.e., a genetic test for Alzheimer's risk gene), medical record review, physician questionnaires, for case ascertainment. Potential data sources may include AD registries and prospective community-based cohort studies (e.g., the Cache County Study [Tschanz et al., 2005] and the Cardiovascular Health Study [Fitzpatrick et al., 2004]).

Studies suggest that more than 60% of individuals with AD have three or more comorbid conditions [Doraiswamy et al., 2002], and that comorbidities are an important driver of increased expenditures for AD [Gutterman et al., 1999]. We found that individuals with AD who had more comorbidities were more likely to incur high and persistently high expenditures. A more thorough understanding of this uneven distribution could be achieved by examining differences in specific service use and treatment patterns between individuals with persistently high and low expenditures. Which component of the total spending (e.g., hospitalization or outpatient care) is driving the expenditure persistence? Are there disparities by expenditure group in managing comorbid conditions, such as psychiatric disorders and diabetes, after controlling for disease severity? It would be interesting to

explore whether there exists subgroups who had relatively few comorbidities and few functional limitations but utilize a large share of medical services, and discuss opportunities for health plans to improve care coordination and lower cost. In addition, large claims databases from MCOs, such as the Integrated Healthcare Information Services (IHCIS), can be used to validate our analysis of expenditure concentration and persistence, and to evaluate specific utilization differences in inpatient care, outpatient visits, and prescription drugs between individuals with persistently high and low expenditures.

In the comparison analysis of various risk adjustment measures, functional ability as measured by ADLs appeared to be a promising frailty adjuster to the diagnosis-based CMS-HCC model in the AD population. In the future, it will be important to validate self- and proxy-reported functional status using alternative and possibly more objective measures, such as physician rating or having people perform the task during survey interview, to improve predictive accuracy. Using measures specifically for assessing daily living abilities among individuals with AD, such as the Bristol Activities of Daily Living Scale [Bucks et al., 1996], may improve predictive power, but disease-specific measures may not be applicable to the overall Medicare population. Moreover, future research should assess the time and cost of collecting additional information as functional status is not routinely available in administrative data, and the acquisition, cleaning and use of both administrative and survey data are often time-consuming and costly [Powers et al., 2005].

In the MCBS, prescription drug expenditures were imputed data aggregated from all payments but not actual expenditure data. Future work is needed to validate the current study using actual drug expenditures, such as large claims database from MCOs. With available pharmacy claims data, an extension of the current study would be to compare the performance of pharmacy-based risk adjustment measures against other measures in predicting overall and drug expenditures, and to discuss whether these measures provide a viable alternative approach to improving prediction for individuals with substantial disabilities,

such as the AD population. Beginning 2006, outpatient prescription drug coverage was provided to all Medicare beneficiaries through Part D benefits. Risk adjustment for Medicare prescription drug plan payments will require review and update regularly as older drugs are going generic and newer, pricy drugs are becoming available [Robst et al., 2007].

Additional research should expand the scope to include long-term-care residents in that they may have different utilization patterns from community-dwelling individuals with AD. In long-term care, research has shown that the specific underlying condition is less important for measuring risk, whereas the extent of functional limitations may have a significant role for predicting both clinical and economic outcomes [Berlowitz and Rosen, 2003]. Diagnosis-based models with frailty adjustment will require further review and refinement for risk adjustment to be applicable to long-term-care settings. For example, an admission profile, such as the pressure ulcer prevalence rate, may be adjusted to recognize that some facilities admit sicker individuals [Berlowitz and Rosen, 2003]. Future investigation should determine how to scale functional limitations (e.g., count of functional impairments or dummy variables for specific impairment) and how to assign weights to these alternative frailty adjusters.

REFERENCES

Alzheimer's Association. 2001. "Medicare and Medicaid Costs for People with Alzheimer's Disease."

Alzheimer's Association. 2007. "Alzheimer's Disease Facts and Figures." Chicago, IL,

Ash, A. S., Ellis, R. P., Pope, G. C., Ayanian, J. Z., Bates, D. W., Burstin, H., Iezzoni, L. I., MacKay, E. and Yu, W. 2000. "Using diagnoses to describe populations and predict costs." *Health Care Financ Rev* 21 (3): 7-28.

Ash, A. S., Zhao, Y., Ellis, R. P. and Schlein Kramer, M. 2001. "Finding future high-cost cases: comparing prior cost versus diagnosis-based methods." *Health Serv Res* 36 (6 Pt 2): 194-206.

Banthin, J. S. and Miller, E. 2006. "Persistence in Medicare Prescription Drug Expenditures by Treatment Class." Agency for Healthcare Research and Quality Working Paper No. 06006, http://gold.ahrq.gov.

Barnett, A. G., van der Pols, J. C. and Dobson, A. J. 2005. "Regression to the mean: what it is and how to deal with it." *Int J Epidemiol* 34 (1): 215-20.

Beebe, J. C. 1988. "Medicare reimbursement and regression to the mean." *Health Care Financ Rev* 9 (3): 9-22.

Berk, M. L. and Monheit, A. C. 2001. "The concentration of health care expenditures, revisited." *Health Aff (Millwood)* 20 (2): 9-18.

Berlowitz, D. and Rosen, A. K. 2003. "Ch16: Risk adjustment for studying long-term-care." Risk adjustment for measuring health care outcomes, 3rd ed., Chicago, Illinois, Health Administration Press.

Bloom, B. S., de Pouvourville, N. and Straus, W. L. 2003. "Cost of illness of Alzheimer's disease: how useful are current estimates?" *Gerontologist* 43 (2): 158-64.

Blough, D. K., Madden, C. W. and Hornbrook, M. C. 1999. "Modeling risk using generalized linear models." *J Health Econ* 18 (2): 153-71.

Brookmeyer, R., Gray, S. and Kawas, C. 1998. "Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset." *Am J Public Health* 88 (9): 1337-42.

Brummel-Smith, K. 2001. "How much does it really cost to care for persons with Alzheimer's disease?" *J Am Geriatr Soc* 49 (9): 1255-6.

Bucks, R. S., Ashworth, D. L., Wilcock, G. K. and Siegfried, K. 1996. "Assessment of activities of daily living in dementia: development of the Bristol Activities of Daily Living Scale." *Age Ageing* 25 (2): 113-20.

Buntin, M. B. and Zaslavsky, A. M. 2004. "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures." *J Health Econ* 23 (3): 525-42.

Bureau of Labor Statistics. 2007. U.S. Department of Labor. Accessed on September 15, 2007. Available at: http://www.bls.gov/cpi/home.htm

Caro, J. J., Getsios, D., Migliaccio-Walle, K., Raggio, G. and Ward, A. 2001. "Assessment of health economics in Alzheimer's disease (AHEAD) based on need for full-time care." *Neurology* 57 (6): 964-71.

Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, C. R. 1987. "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." *J Chronic Dis* 40 (5): 373-83.

CMS. 2003. "Appendix A: Technical Documentation for the Medicare Current Beneficiary Survey." Accessed on October 10, 2007. Available at: http://www.cms.hhs.gov/mcbs/downloads/HHC2003appendixA.pdf

CMS. 2006a. "Risk adjustment model software." Accessed on December 15, 2007. Available at: http://www.cms.hhs.gov/MedicareAdvtgSpecRateStats/06_Risk_adjustment.asp#TopOfPageCMS Home > Research, Statistics, Data and Systems > Medicare Advantage - Rates & Statistics > Risk adjustment

CMS. 2006b. "Survey Overview." Accessed on November 1, 2006. Available at: http://www.cms.hhs.gov/apps/mcbs/Overview.asp

CMS. 2006c. "Medicare managed care manual." Accessed on May 1, 2007. Available at: http://www.cms.hhs.gov/Manuals/IOM/list.asp

CMS. 2007a. "Medicare Current Beneficiary Survey CY 2004 Cost and Use, Section 5: Filling the Gaps."

CMS. 2007b. "National Health Expenditure Accounts: 2006 Highlights." Accessed on April 8, 2008. Available at: http://www.cms.hhs.gov/nationalhealthexpenddata/downloads/highlights.pdf CMS Home > Research, Statistics, Data and Systems > National Health Expenditure Data > Historical

CMS. 2008. "Nursing Home Quality Initiatives." Accessed on March 17, 2008. Available at: http://www.cms.hhs.gov/NursingHomeQualityInits/01_Overview.asp#TopOfPage CMS Home > Medicare > Nursing Home Quality Initiatives

Cohen, J. 1960. "A coefficient of agreement for nominal scales." *Educational and Psychological Measurement* 20 37-46.

Coulson, N. E. and Stuart, B. 1992. "Persistence in the use of pharmaceuticals by the elderly. Evidence from annual claims." *J Health Econ* 11 (3): 315-28.

Cucciare, M. A. and O'Donohue, W. 2006. "Predicting future healthcare costs: how well does risk-adjustment work?" *J Health Organ Manag* 20 (2-3): 150-62.

Cumming, R. B., Knutsin, D., Cameron, B. A. and Derrick, B. 2002. "A comparative analysis of claims-based methods of health risk assessment for commercial
populations. Society of Actuaries." Accessed on February 15, 2007. Available at: http://www.soa.org/research/files/pdf/_asset_id=2583046.pdf

D'Hoore, W., Sicotte, C. and Tilquin, C. 1993. "Risk adjustment in outcome assessment: the Charlson comorbidity index." *Methods Inf Med* 32 (5): 382-7.

DeBettignies, B. H., Mahurin, R. K. and Pirozzolo, F. J. 1990. "Insight for impairment in independent living skills in Alzheimer's disease and multi-infarct dementia." *J Clin Exp Neuropsychol* 12 (2): 355-63.

Deyo, R. A., Cherkin, D. C. and Ciol, M. A. 1992. "Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases." *J Clin Epidemiol* 45 (6): 613-9.

Doraiswamy, P. M., Leon, J., Cummings, J. L., Marin, D. and Neumann, P. J. 2002. "Prevalence and impact of medical comorbidity in Alzheimer's disease." *J Gerontol A Biol Sci Med Sci* 57 (3): M173-7.

Duan, N., Manning, W. G., Morris, C. N. and Newhouse, J. P. 1983. "A comparison of alternative models for the demand for medical care." *Journal of Business & Economic Statistics* 1 (2): 115-26.

Ellis, R. P., Pope, G. C., Iezzoni, L., Ayanian, J. Z., Bates, D. W., Burstin, H. and Ash, A. S. 1996. "Diagnosis-based risk adjustment for Medicare capitation payments." *Health Care Financ Rev* 17 (3): 101-28.

England, A., Hubbell, K., Judkins, D. and Ryaboy, S. 1994. "Imputation of medical cost and payment data." *Proceedings of the Section on Survey Research Methods, American Statistical Association* 406-11.

Epstein, A. M. and Cumella, E. J. 1988. "Capitation payment: using predictors for medical utilization to adjust rates." *Health Care Financ Rev* 10 (1): 51-69.

Evans, D. A., Funkenstein, H. H., Albert, M. S., Scherr, P. A., Cook, N. R., Chown, M. J., Hebert, L. E., Hennekens, C. H. and Taylor, J. O. 1989. "Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported." *JAMA* 262 (18): 2551-6.

Evans, D. A. 1990. "Estimated prevalence of Alzheimer's disease in the United States." *Milbank Q* 68 (2): 267-89.

Farias, S. T., Mungas, D. and Jagust, W. 2005. "Degree of discrepancy between self and other-reported everyday functioning by cognitive status: dementia, mild cognitive impairment, and healthy elders." *Int J Geriatr Psychiatry* 20 (9): 827-34.

Farley, J. F., Harley, C. R. and Devine, J. W. 2006. "A comparison of comorbidity measurements to predict healthcare expenditures." *Am J Manag Care* 12 (2): 110-9.

Fillit, H., Geldmacher, D. S., Welter, R. T., Maslow, K. and Fraser, M. 2002a. "Optimizing coding and reimbursement to improve management of Alzheimer's disease and related dementias." *J Am Geriatr Soc* 50 (11): 1871-8.

Fillit, H., Hill, J. W. and Futterman, R. 2002b. "Health care utilization and costs of Alzheimer's disease: the role of co-morbid conditions, disease stage, and pharmacotherapy." *Fam Med* 34 (7): 528-35.

Fillit, H. and Hill, J. 2004. "The economic benefits of acetylcholinesterase inhibitors for patients with Alzheimer's disease and associated dementias." *Alzheimer Dis Assoc Disord* 18 Suppl 1 S24-9.

Fillit, H. M. 2000. "The pharmacoeconomics of Alzheimer's disease." *Am J Manag Care* 6 (22 Suppl): S1139-44; discussion S45-8.

Fishman, P. A., Goodman, M. J., Hornbrook, M. C., Meenan, R. T., Bachman, D. J. and O'Keeffe Rosetti, M. C. 2003. "Risk adjustment using automated ambulatory pharmacy data: the RxRisk model." *Med Care* 41 (1): 84-99.

Fitzpatrick, A. L., Kuller, L. H., Ives, D. G., Lopez, O. L., Jagust, W., Breitner, J. C., Jones, B., Lyketsos, C. and Dulberg, C. 2004. "Incidence and prevalence of dementia in the Cardiovascular Health Study." *J Am Geriatr Soc* 52 (2): 195-204.

Fleishman, J. A., Cohen, J. W., Manning, W. G. and Kosinski, M. 2006. "Using the SF-12 health status measure to improve predictions of medical expenditures." *Med Care* 44 (5 Suppl): I54-63.

Folstein, M. F., Folstein, S. E. and McHugh, P. R. 1975. ""Mini-mental state". A practical method for grading the cognitive state of patients for the clinician." *J Psychiatr Res* 12 (3): 189-98.

Fowles, J. B., Weiner, J. P., Knutson, D., Fowler, E., Tucker, A. M. and Ireland, M. 1996. "Taking health status into account when setting capitation rates: a comparison of risk-adjustment methods." *JAMA* 276 (16): 1316-21.

Ganguli, M., Dodge, H. H., Shen, C., Pandav, R. S. and DeKosky, S. T. 2005. "Alzheimer's disease and mortality: a 15-year epidemiological study." *Arch Neurol* 62 (5): 779-84.

Garber, A. M., MaCurdy, T. E. and McClellan, M. C. 1997. "Persistence of Medicare Expenditures Among Elderly Beneficiaries." NBER Working Paper Series. National Bureau of Economic Research. Accessed on Dec, 12, 2006. Available at: http://www.nber.org/papers/w6249

Ghali, W. A., Hall, R. E., Rosen, A. K., Ash, A. S. and Moskowitz, M. A. 1996. "Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data." *J Clin Epidemiol* 49 (3): 273-8.

Gorina, Y., Hoyert, D., Lentzner, H. and Goulding, M. 2006. "Trends in causes of death among older persons in the United States." Aging Trends, No 6. Hyattsville, Maryland, National Center for Health Statistics.

Greenwald, L. M. 2000. "Medicare risk-adjusted capitation payments: from research to implementation." *Health Care Financ Rev* 21 (3): 1-5.

Grossberg, G. T. 2002. "The ABC of Alzheimer's disease: behavioral symptoms and their treatment." *Int Psychogeriatr* 14 Suppl 1 27-49.

Gutterman, E. M., Markowitz, J. S., Lewis, B. and Fillit, H. 1999. "Cost of Alzheimer's disease and related dementia in managed-medicare." *J Am Geriatr Soc* 47 (9): 1065-71.

Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A. and Evans, D. A. 2003. "Alzheimer disease in the US population: prevalence estimates using the 2000 census." *Arch Neurol* 60 (8): 1119-22.

Hill, J., Fillit, H., Thomas, S. K. and Chang, S. 2006. "Functional impairment, healthcare costs and the prevalence of institutionalisation in patients with Alzheimer's disease and other dementias." *PharmacoEconomics* 24 (3): 265-80.

Hill, J. W., Futterman, R., Duttagupta, S., Mastey, V., Lloyd, J. R. and Fillit, H. 2002. "Alzheimer's disease and related dementias increase costs of comorbidities in managed Medicare." *Neurology* 58 (1): 62-70.

Hughes, J. S., Averill, R. F., Eisenhandler, J., Goldfield, N. I., Muldoon, J., Neff, J. M. and Gay, J. C. 2004. "Clinical Risk Groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management." *Med Care* 42 (1): 81-90.

Iezzoni, L. I. 2003a. "Ch.3: Range of risk factors." Risk adjustment for measuring health care outcomes, 3rd ed., Chicago, Illinois, Health Administration Press.

Iezzoni, L. I. 2003b. "Ch.1: Reasons for risk adjustment." Risk adjustment for measuring health care outcomes, 3rd ed., Chicago, Illinois, Health Administration Press.

Johnson, N., Davis, T. and Bosanquet, N. 2000. "The epidemic of Alzheimer's disease. How can we manage the costs?" *PharmacoEconomics* 18 (3): 215-23.

Kane, R. L. and Atherly, A. 2000. "Medicare expenditures associated with Alzheimer disease." *Alzheimer Dis Assoc Disord* 14 (4): 187-95.

Kapur, K., Young, A. S. and Murata, D. 2000. "Risk adjustment for high utilizers of public mental health care." *J Ment Health Policy Econ* 3 (3): 129-37.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A. and Jaffe, M. W. 1963. "Studies of Illness in the Aged. the Index of ADL: a Standardized Measure of Biological and Psychosocial Function." *JAMA* 185 914-9.

Katzman, R. 1986. "Alzheimer's disease." *N Engl J Med* 314 (15): 964-73.

Kaufer, D. I., Borson, S., Kershaw, P. and Sadik, K. 2005. "Reduction of caregiver burden in Alzheimer's disease by treatment with galantamine." *CNS Spectr* 10 (6): 481-8.

Kautter, J. and Pope, G. C. 2004. "CMS frailty adjustment model." *Health Care Financing Review* 26 (2): 1-19.

Koss, E., Patterson, M. B., Ownby, R., Stuckey, J. C. and Whitehouse, P. J. 1993. "Memory evaluation in Alzheimer's disease. Caregivers' appraisals and objective testing." *Arch Neurol* 50 (1): 92-7.

Kronick, R., Dreyfus, T., Lee, L. and Zhou, Z. Y. 1996. "Diagnostic risk adjustment for Medicaid: The disability payment system." *Health Care Financing Review* 17 (3): 7-33.

Kronick, R., Gilmer, T., Dreyfus, T. and Lee, L. 2000. "Improving health-based payment for Medicaid beneficiaries: CDPS." *Health Care Financ Rev* 21 (3): 29-64.

Kronick, R., Gilmer, T., Dreyfus, T. and Ganiats, T. G. 2002. "CDPS-Medicare: The Chronic Illness and Disability Payment System Modified to Predict Expenditures for Medicare Beneficiaries." Accessed on January 25, 2008. Available at: http://cdps.ucsd.edu/CDPS_Medicare.pdf

Lamers, L. M. 1999. "Risk-adjusted capitation based on the Diagnostic Cost Group Model: an empirical evaluation with health survey information." *Health Serv Res* 33 (6): 1727-44.

Landis, J. R. and Koch, G. G. 1977. "The measurement of observer agreement for categorical data." *Biometrics* 33 (1): 159-74.

Larson, E. B., Shadlen, M. F., Wang, L., McCormick, W. C., Bowen, J. D., Teri, L. and Kukull, W. A. 2004. "Survival after initial diagnosis of Alzheimer disease." *Ann Intern Med* 140 (7): 501-9.

Leon, J., Cheng, C. K. and Neumann, P. J. 1998. "Alzheimer's disease care: costs and potential savings." *Health Aff (Millwood)* 17 (6): 206-16.

Leung, G. M., Yeung, R. Y., Chi, I. and Chu, L. W. 2003. "The economics of Alzheimer disease." *Dement Geriatr Cogn Disord* 15 (1): 34-43.

Lin, P. J. 2008a. "Identifying community-dwelling individuals with Alzheimer's disease from observational data." Doctoral dissertation. Chapel Hill, NC, UNC-CH.

Lin, P. J. 2008b. "The concentration and persistence of health care expenditures and prescription drug expenditures in Medicare beneficiaries with Alzheimer's disease." Doctoral dissertation. Chapel Hill, NC, UNC-CH.

Liptak, G. S., Shone, L. P., Auinger, P., Dick, A. W., Ryan, S. A. and Szilagyi, P. G. 2006. "Short-term persistence of high health care costs in a nationally representative sample of children." *Pediatrics* 118 (4): e1001-9.

Long, K., Sudha, S. and Mutran, E. J. 1998. "Elder-proxy agreement concerning the functional status and medical history of the older person: the impact of caregiver burden and depressive symptomatology." *J Am Geriatr Soc* 46 (9): 1103-11.

Maciejewski, M. L., Liu, C. F., Derleth, A., McDonell, M., Anderson, S. and Fihn, S. D. 2005. "The performance of administrative and self-reported measures for risk adjustment of Veterans Affairs expenditures." *Health Serv Res* 40 (3): 887-904.

Maddala, G. S. 1983. "Limited-Dependent and Qualitative Variables in Econometrics." Cambridge, Cambridge University Press.

Manning, W. G. and Mullahy, J. 2001. "Estimating log models: to transform or not to transform?" *J Health Econ* 20 (4): 461-94.

McNemar, Q. 1947. "Note on the sampling error of the difference between correlated proportions or percentages." *Psychometrika* 12 153-57.

Meenan, R. T., O'Keeffe-Rosetti, C., Hornbrook, M. C., Bachman, D. J., Goodman, M. J., Fishman, P. A. and Hurtado, A. V. 1999. "The sensitivity and specificity of forecasting high-cost users of medical care." *Med Care* 37 (8): 815-23.

Meenan, R. T., Goodman, M. J., Fishman, P. A., Hornbrook, M. C., O'Keeffe-Rosetti, M. C. and Bachman, D. J. 2003. "Using risk-adjustment models to identify high-cost risks." *Med Care* 41 (11): 1301-12.

Menzin, J., Lang, K., Friedman, M., Neumann, P. and Cummings, J. L. 1999. "The economic cost of Alzheimer's disease and related dementias to the California Medicaid program ("Medi-Cal") in 1995." *Am J Geriatr Psychiatry* 7 (4): 300-8.

Mohs, R. C., Schmeidler, J. and Aryan, M. 2000. "Longitudinal studies of cognitive, functional and behavioural change in patients with Alzheimer's disease." *Stat Med* 19 (11-12): 1401-9.

Monheit, A. C. 2003. "Persistence in health expenditures in the short run: prevalence and consequences." *Med Care* 41 (7 Suppl): III53-III64.

Needham, D. M., Scales, D. C., Laupacis, A. and Pronovost, P. J. 2005. "A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research." *J Crit Care* 20 (1): 12-9.

Neumann, P. J., Araki, S. S. and Gutterman, E. M. 2000. "The use of proxy respondents in studies of older adults: lessons, challenges, and opportunities." *J Am Geriatr Soc* 48 (12): 1646-54.

Newcomer, R., Clay, T., Luxenberg, J. S. and Miller, R. H. 1999. "Misclassification and selection bias when identifying Alzheimer's disease solely from Medicare claims records." *J Am Geriatr Soc* 47 (2): 215-9.

Newhouse, J. P., Manning, W. G., Keeler, E. B. and Sloss, E. M. 1989. "Adjusting capitation rates using objective health measures and prior utilization." *Health Care Financ Rev* 10 (3): 41-54.

Newhouse, J. P. 1996. "Reimbursing health plans and health providers: Efficiency in production versus selection." *Journal of Economic Literature* 34 (3): 1236-63.

Newhouse, J. P., Buntin, M. B. and Chapman, J. D. 1997. "Risk adjustment and Medicare: taking a closer look." *Health Aff (Millwood)* 16 (5): 26-43.

Noyes, K., Liu, H. and Temkin-Greener, H. 2006. "Cost of caring for Medicare beneficiaries with Parkinson's disease: Impact of the CMS-HCC risk-adjustment model." *Dis Manag* 9 (6): 339-48.

Pacala, J. T., Boult, C., Urdangarin, C. and McCaffrey, D. 2003. "Using self-reported data to predict expenditures for the health care of older people." *J Am Geriatr Soc* 51 (5): 609-14.

Pope, G. C., Adamache, K. W., Walsh, E. G. and Khandker, R. K. 1998. "Evaluating alternative risk adjusters for Medicare." *Health Care Financing Review* 20 (2): 109.

Pope, G. C., Ellis, R. P., Ash, A. S., Liu, C. F., Ayanian, J. Z., Bates, D. W., Burstin, H., Iezzoni, L. I. and Ingber, M. J. 2000. "Principal inpatient diagnostic cost group model for Medicare risk adjustment." *Health Care Financ Rev* 21 (3): 93-118.

Pope, G. C., Kautter, J., Ellis, R. P., Ash, A. S., Ayanian, J. Z., Lezzoni, L. I., Ingber, M. J., Levy, J. M. and Robst, J. 2004. "Risk adjustment of Medicare capitation payments using the CMS-HCC model." *Health Care Financ Rev* 25 (4): 119-41.

Powers, C. A., Meyer, C. M., Roebuck, M. C. and Vaziri, B. 2005. "Predictive modeling of total healthcare costs using pharmacy claims data: a comparison of alternative econometric cost modeling techniques." *Med Care* 43 (11): 1065-72.

Pressley, J. C., Trott, C., Tang, M., Durkin, M. and Stern, Y. 2003. "Dementia in community-dwelling elderly patients: A comparison of survey data, Medicare claims, cognitive screening, reported symptoms, and activity limitations." *J Clin Epidemiol* 56 (9): 896-905.

Quan, H., Parsons, G. A. and Ghali, W. A. 2002. "Validity of information on comorbidity derived rom ICD-9-CCM administrative data." *Med Care* 40 (8): 675-85.

Rice, D. P., Fillit, H. M., Max, W., Knopman, D. S., Lloyd, J. R. and Duttagupta, S. 2001. "Prevalence, costs, and treatment of Alzheimer's disease and related dementia: a managed care perspective." *Am J Manag Care* 7 (8): 809-18.

Riley, G. F. 2000. "Risk adjustment for health plans disproportionately enrolling frail Medicare beneficiaries." *Health Care Financ Rev* 21 (3): 135-48.

Robinson, J. and Karon, S. L. 2000. "Modeling Medicare costs of PACE (Program of All-Inclusive Care for the Elderly) populations." *Health Care Financ Rev* 21 (3): 149-70.

Robst, J., Levy, J. M. and Ingber, M. J. 2007. "Diagnosis-based risk adjustment for Medicare prescription drug plan payments." *Health Care Financ Rev* 28 (4): 15-30.

Romano, P. S., Roos, L. L. and Jollis, J. G. 1993. "Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives." *J Clin Epidemiol* 46 (10): 1075-9; discussion 81-90.

Rosen, W. G., Mohs, R. C. and Davis, K. L. 1984. "A new rating scale for Alzheimer's disease." *Am J Psychiatry* 141 (11): 1356-64.

Rothman, M. L., Hedrick, S. C., Bulcroft, K. A., Hickam, D. H. and Rubenstein, L. Z. 1991. "The validity of proxy-generated scores as measures of patient health status." *Med Care* 29 (2): 115-24.

Russell, L. B. and Chaudhuri, A. 1992. "The inequality of medical expenditures for several years in a healthy, nonelderly population." *Med Care* 30 (10): 908-16.

Schneeweiss, S. and Maclure, M. 2000. "Use of comorbidity scores for control of confounding in studies using administrative databases." *Int J Epidemiol* 29 (5): 891-8.

Shwartz, M. and Ash, A. 2003. "Ch.10: Evaluating risk-adjustment models empirically." Risk adjustment for measuring health care outcomes, 3rd ed., Chicago, Illinois, Health Administration Press.

Simoni-Wastila, L., Shaffer, T. and Stuart, B. 2007. "A National Comparison of Prescription Drug Expenditures by Medicare Beneficiaries Living in the Community and Long-Term Care Facility Settings." U.S. Department of Health and Human Services. Accessed on January 18, 2008. Available at: http://aspe.hhs.gov/daltcp/reports/2007/pdnatcom.htm

Sloane, P. D., Zimmerman, S., Suchindran, C., Reed, P., Wang, L., Boustani, M. and Sudha, S. 2002. "The public health impact of Alzheimer's disease, 2000-2050: potential implication of treatment advances." *Annu Rev Public Health* 23 213-31.

Stuart, B., Ahern, F., Rabatin, V. and Johnson, A. 1991. "Patterns of outpatient prescription drug use among Pennsylvania elderly." *Health Care Financ Rev* 12 (3): 61-72.

Szklo, M. and Nieto, F. J. 2004. "Epidemiology: Beyond the Basics." Salisbury, MA, Jones and Bartlett Publishers.

Taylor, D. H., Jr. and Sloan, F. A. 2000. "How much do persons with Alzheimer's disease cost Medicare?" *J Am Geriatr Soc* 48 (6): 639-46.

Taylor, D. H., Jr., Schenkman, M., Zhou, J. and Sloan, F. A. 2001. "The relative effect of Alzheimer's disease and related dementias, disability, and comorbidities on cost of care for elderly persons." *J Gerontol B Psychol Sci Soc Sci* 56 (5): S285-93.

Temkin-Greener, H., Meiners, M. R. and Gruenberg, L. 2001. "PACE and the Medicare+Choice risk-adjusted payment model." *Inquiry* 38 (1): 60-72.

Tierney, M. C., Szalai, J. P., Snow, W. G. and Fisher, R. H. 1996. "The prediction of Alzheimer disease. The role of patient and informant perceptions of cognitive deficits." *Arch Neurol* 53 (5): 423-7.

Tschanz, J. T., Treiber, K., Norton, M. C., Welsh-Bohmer, K. A., Toone, L., Zandi, P. P., Szekely, C. A., Lyketsos, C. and Breitner, J. C. 2005. "A population study of Alzheimer's disease: findings from the Cache County Study on Memory, Health, and Aging." *Care Manag J* 6 (2): 107-14.

US GAO. 1998. "Alzheimer's Disease: Estimates of Prevalence in the US." Washington, DC, US General Accounting Office.

Von Korff, M., Wagner, E. H. and Saunders, K. 1992. "A chronic disease score from automated pharmacy data." *J Clin Epidemiol* 45 (2): 197-203.

Weiner, M., Powe, N. R., Weller, W. E., Shaffer, T. J. and Anderson, G. F. 1998. "Alzheimer's disease under managed care: implications from Medicare utilization and expenditure patterns." *J Am Geriatr Soc* 46 (6): 762-70.

Wooldridge, J. M. 2003. "Introductory Econometrics: A Modern Appraoch." South-Western, Mason, OH.

Wrobel, M. V., Doshi, J., Stuart, B. C. and Briesacher, B. 2003. "Predictability of prescription drug expenditures for Medicare beneficiaries." *Health Care Financing Review* 25 (2): 46.

Zhao, Y., Ellis, R. P., Ash, A. S., Calabrese, D., Ayanian, J. Z., Slaughter, J. P., Weyuker, L. and Bowen, B. 2001. "Measuring population health risks using inpatient diagnoses and outpatient pharmacy data." *Health Serv Res* 36 (6 Pt 2): 180-93.

Zhao, Y., Ash, A. S., Ellis, R. P., Ayanian, J. Z., Pope, G. C., Bowen, B. and Weyuker, L. 2005. "Predicting pharmacy costs and other medical costs using diagnoses and drug claims." *Med Care* 43 (1): 34-43.