



Published in final edited form as:

Alzheimer Dis Assoc Disord. 2015 ; 29(1): 75–81. doi:10.1097/WAD.0000000000000048.

Errors in Self-Reports of Health Services Use: Impact on Alzheimer Disease Clinical Trial Designs

Christopher M. Callahan, MD^{a,b,e}, Wanzhu Tu, PhD^{a,c,e}, Timothy E. Stump, MA^c, Daniel O. Clark, PhD^{a,b,e}, Kathleen Unroe, MD, MHA^{a,b,e}, and Hugh C. Hendrie, MB, ChB, DSc^{a,d,e}

^aIndiana University Center for Aging Research Indianapolis, Indiana, USA

^bDepartment of Medicine, Indiana University School of Medicine, Indianapolis, Indiana, USA

^cDepartment of Biostatistics, Indiana University School of Medicine, Indianapolis, Indiana, USA

^dDepartment of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana, USA

^eRegenstrief Institute, Inc., Indianapolis, Indiana, USA

Abstract

Background—Most Alzheimer's disease clinical trials that compare the use of health services rely on reports of caregivers. The goal of this study was to assess the accuracy of self-reports among older adults with Alzheimer's disease and their caregiver proxy respondents. This issue is particularly relevance to Alzheimer's disease clinical trials because inaccuracy can lead both to loss of power and increased bias in study outcomes.

Methods—We compared respondent accuracy in reporting any use and in reporting the frequency of use with actual utilization data as documented in a comprehensive database. We next simulated the impact of under-reporting and over-reporting on sample size estimates and treatment effect bias for clinical trials comparing utilization between experimental groups.

Results—Respondents self-reports have a poor level of accuracy with Kappa values often below 0.5. Respondents tend to under-report use even for rare events such as hospitalizations and nursing home stays. In analyses simulating under- and over-reporting of varying magnitude, we found that errors in self-reports can increase the required sample size by 15–30%. In addition, bias in the reported treatment effect ranged from 3%–18% due to both under-reporting and over-reporting errors.

Conclusions—Use of self-report data in clinical trials of Alzheimer's disease treatments may inflate sample size needs. . Even when adequate power is achieved by increasing sample size, reporting errors can result in a biased estimate of the true effect size of the intervention.

Keywords

Alzheimer's disease; Self-reports; Health care utilization; Clinical trials; Health services research

INTRODUCTION

Clinical trials investigating the effectiveness of new treatments for Alzheimer's disease may report outcomes across multiple domains including outcomes relevant to the patient, the caregiver, and the health care system.¹⁻⁷ In the typical framework explaining the relationship between these outcomes, scientists might hypothesize that an intervention improves patient cognition or behavior which then leads to improved patient quality of life and reduced caregiver stress and these improvements thereby lead to decreased health care use.⁸⁻¹⁰ Among older adults with Alzheimer's disease, reducing health care cost is an important target because older adults with dementia accrue much higher health care costs than those without dementia.¹¹⁻¹³ Delaying time to institutionalization has been a particularly high-value goal because this is an important outcome for patients, caregivers, and third-party payers.^{8,9,14} Indeed, nursing home care is so expensive that treatments that might be considered otherwise too expensive can be justified if they are able to demonstrate savings in the cost of nursing facility care.^{8,15-17}

Resource use surveys have been developed for use in general population studies and also for use among older adults with dementia.¹⁸⁻²⁰ These instruments query respondents and their caregivers about the frequency of episodes of care, lengths of stay for hospitalizations or nursing facility care, and may also seek to identify specific dates, for example, of institutionalization. Some of these surveys also seek to quantify the cost of unpaid caregiving. Most Alzheimer's disease clinical trials that measure the use of health services rely on surveys of caregiver's self-reports of patient's use of health services.^{3,18,21-23} The "proxy self-reports" of health care use can be converted to estimates of total health care costs by assigning an average monetary value to each episode of care, such as the average cost of a Medicare hospital stay.

Self-reports of health service use are known to have substantial error.²⁴⁻²⁸ Respondent errors may be due to recall bias, telescoping (recall of events correctly but not in the correct time frame), or perceptions of social acceptability.²⁸ The potential bias introduced by systematic and non-systematic error in self-reported health care use varies based on the study population, hypothesis, and design.²⁸ We previously conducted a randomized controlled clinical trial of collaborative care for older adults with Alzheimer's disease.²⁹ As part of this clinical trial, we collected data on resource use from subjects' caregivers using the Resource Use Inventory.³⁰ We also obtained utilization data from the local electronic medical record to supplement the self-reported data in the original clinical trial. More recently, we obtained subjects' actual Medicare and Medicaid claims data as well as their standardized nursing home and home health care assessments. The specific aim of this study was to assess the accuracy of self-reports of health care use among older adults with Alzheimer's disease and their caregiver proxy respondents. We then use these findings to simulate the impact of errors in self-reports on sample size and power considerations in clinical trials of Alzheimer's disease seeking to demonstrate difference in health care use between experimental groups. We pursued these simulations because most Alzheimer's disease clinical trials are small (<200 subjects) and yet the consumers of these studies often demand cost-effectiveness evidence which is most often derived through data on self-reported utilization. Thus, error within these self-reports could result in study conclusions

that show no effect of the intervention on utilization when one actually exists (due to inadequate statistical power). Even in the setting of adequate power, the study could report an effect of the intervention that was biased (an effect size that is erroneously too high or too low) due to errors from self-reports.

METHODS

This study was approved by the Indiana University Purdue University- Indianapolis Institutional Review Board and the Centers for Medicare and Medicaid Services Privacy Board. The current report focuses on the 100 subjects enrolled from an urban safety net health system serving the Indianapolis area and randomized to intervention or control groups.²⁹ Because the study intervention did not alter capacity for self-reported health care use, we report the results of these 100 subjects in aggregate. All subjects enrolled in the original trial were enrolled as dyads so that all subjects with Alzheimer's disease were enrolled with an informal caregiver. The original clinical trial was powered to demonstrate clinically significant differences in the Neuropsychiatric Inventory and was not powered to demonstrate differences in health care use between experimental groups.

Data from this study come from two sources: (1) self-reported utilization data as reported by the caregiver and collected in the prior clinical trial; and (2) actual utilization data as documented in medical records and claims data for the same period covered by the self-reported utilization data.¹¹ The clinical trial used the Alzheimer's Disease Cooperative Studies Group Resource Use Inventory.^{19,30} This instrument was delivered by professional research assistants in telephone interviews at baseline (covering the 6 months prior to baseline), 6, 12, and 18 months. The Resource Use Inventory includes questions about assistive devices, examinations by doctors or nurses, use of adult day care, and hours of non-professional caregiving. For the purposes of this project we focused on the four items below:

- In the last six months, how many different visits has [subject] had to a hospital emergency room or urgent care facility
- In the last six months, how many times was [subject] admitted to the hospital?
- In the last six months, how many nights did [subject] spend in the nursing home?
- In the last 6 months, how many days did a home health aide, attendant or any other paid individual look after [subject]?

We focused on these four outcomes because they represent the high cost health care utilization patterns of concern to third-party payers and the utilization patterns that Alzheimer's disease interventions seek to decrease. Actual utilization data were documented in a unique dataset in which we merged the clinical data available from a regional health information exchange³¹ with four additional databases: (1) Medicare claims; (2) Resident Level Minimum Data Set (MDS); (3) Outcome and Assessment Information Set; (OASIS) and (4) Indiana Medicaid claims. Through this comprehensive database, we can document health care use across health care systems and across the continuum of care. Thus, the data on utilization are not limited to data collected at a single health care system or a single payer. Episodes of care within this database include documentation of actual dates of use so

we can place use within the baseline (–6 months to 0 months), 6 (0–6 months), 12 (7–12 months), and 18 (13–18 months) month windows used in the self-report data. Alternatively, total use over the entire time period can be summed from self-report or claims data. For each patient, we determined all emergency department use, all hospitalizations, all nursing facility use, and all home health use for each of the time periods corresponding to the self-report survey. We used data available in any of the merged datasets to capture any of these episodes of care including MDS and OASIS data.

For each subject, we compared the self-reported data for use in any of the four care sites (emergency department, hospital, nursing home care, or home health care) first by looking at agreement on whether any use occurred. We report kappa statistics for each 6 month period for each of the four sites of care using any use as the comparison unit. This requires the respondent only to correctly report if use in the site occurred in the past six months as a yes or no question. We also report the percentage of patients who over-estimated or under-estimated use. To explore respondents' accuracy on the volume of use, we examined agreement on the volume of use (number of emergency department visits, number of hospitalizations, number of nights in the nursing home, and number of days of home health care).

Using data from the prior clinical trial, we designed a simulation to demonstrate the impact of various rates of under- and over-reporting of health care use on the estimated power for an equal sized two group comparison study. In the simulation, we assumed that a generic self-reported care utilization count for the i th subject followed a negative binomial distribution with mean (i.e., annual rate) of λ_i , and variance $\lambda_i + k \lambda_i^2$, where k was the dispersion parameter indicating the level of data dispersion. Of note, power calculations on utilization counts in most clinical trials are based on a Poisson data assumption. The negative binomial distribution is a more general family of distributions that includes the Poisson distribution as a special case: when $k=0$, $Y_i \sim \text{Poisson}(\lambda_i)$. Under the standard formulation of log linear regression models, we expressed the subject-specific care utilization rate as $\lambda_i = \exp(\beta_0 + \beta_1 \cdot \text{Group}_i + \beta_2 \cdot \text{Under}_i + \beta_3 \cdot \text{Over}_i)$ where β_0 was the logarithmically transformed overall annual rate of utilization; Group_i was the treatment indicator for the i th subject ($\text{Group}_i=1$ if the subject had received the intervention and $\text{Group}_i=0$ if the subject had received usual care), and β_1 was the average treatment effect; Under_i was an indicator for utilization under-reporting ($\text{Under}_i=1$ indicated the utilization count was under-reported, and $\text{Under}_i=0$ if utilization was not under-reported), and β_2 represented the magnitude of under-reporting; similarly, Over_i was an indicator for over-reporting ($\text{Over}_i=1$ if utilization was over-reported), and β_3 represented the magnitude of over-reporting. In the simulation, we set $\beta_0 = \log_e(1) = 0$ (on average one utilization event per year), $\beta_1 = \log_e(0.8)$ (i.e., 20% reduction in utilization rate in intervention group), and for β_2 we considered two different values: $\beta_2 = \log_e(0.7)$ and $\beta_2 = \log_e(0.5)$ (i.e., 30% and 50% reduction in those who under-reported their utilization counts); for β_3 we used $\beta_3 = -\log_e(0.7)$ (i.e., 30% inflation in utilization counts in those who over-reported their utilization counts). We considered scenarios where sample size per treatment arm ranged from 150 to 900 at 150 increment; for a given sample size, we assumed that 1/3 of the subjects who had reported utilization correctly, whereas 1/6, 1/3, or 1/2 of the subjects had

either over-reported or under-reported their care utilization. Using count data generated from the negative binomial model, we compared the simulated power to the situation of no biased-reporting. For each parameter setting, we generated 20,000 data sets. The simulated power corresponding to each sample size was presented graphically.

RESULTS

Table 1 describes the clinical characteristics of the 100 subjects. In general, these subjects are similar to those in typical Alzheimer's disease clinical trials as measured by age, gender, and cognitive function. However, these subjects also suffer from a significant number of comorbid conditions consistent with the original trial's goal to enroll a sample representative of patients with dementia in primary care. Table 2 shows the agreement between actual use and self-reported use when respondents are only required to correctly reported "any use" as compared to "no use". In this table, the cumulative reporting period increases by 6 months across each column. In general, across the four sites of care, respondents tended to under-report use. Even in the setting of rare but hypothetically notable events, such as nursing facility use, there was under-reporting. Table 3 again shows the agreement between actual use and self-reported use when respondents are only required to correctly reported "any use" as compared to "no use". As compared to Table 2, in Table 3 the reporting period is for each 6 month reporting period. In both Tables 2 and 3, kappa statistics are generally poor. Table 4 shows agreement between self-reports of *volume* of use and actual *volume* of use for the first six month reporting period and the first and second six month reporting periods combined (12 month aggregate). Notably, a large percentage of subjects have no utilization in these sites of care and levels of agreement in volume are driven largely by the number of respondents correctly reporting "zero use". Despite the high likelihood of zero use, agreement as measured by the Kappa statistic is poor across all four sites of care.

We then conducted a simulation study to determine the impact of under- and over-reporting on the statistical power of a hypothetical clinical trial of a novel Alzheimer disease intervention seeking to reduce health care costs. We also describe the potential for bias in outcome findings. Herein, we generated count data where intervention subjects had a 20% reduction in utilization rate as compared to control subjects and we contend that a difference of such magnitude would constitute a clinically meaningful and economically important reduction in use. Count data in this scenario could be either hospitalizations or nursing home stays. Although many clinical trials are powered to demonstrate clinically important differences in continuous outcomes such as cognition or behavioral symptoms, we focused specifically on the power implications of amnesic under-reporting on these high-cost utilization outcomes. We considered scenarios of varying proportions of subjects under- and over-reporting their care use as well as the magnitude of under-reporting. In each figure, we report the impact of this error based on proportions of respondents in different scenarios: (1) 1/3 no bias (no errors), 1/3 under-report, 1/3 over-report, (2) 1/3 no bias, 1/2 under-report, 1/6 over-report, and (3) 1/3 no bias, 1/6 under-report and 1/2 over-report. In Figure 1 the magnitude of under-reporting is 30% and in Figure 2 it is 50% (Figure 2) lower than the actual rate of care utilization. In both figures, the magnitude of over-reporting is 30%.

In the absence of self-reporting bias in both Figures 1 and 2, the hypothetical study seeking to demonstrate a 20% reduction in counts of hospitalizations or nursing home stays at 80% power would require approximately 525 subjects per group for a total sample size of 1050. In Figure 1, with 50% of subjects under-reporting events by 30% and 1/6 over-reporting the events by 30%, the required sample size to maintain a power of 80% would grow to approximately 600 subjects per group or a total sample size of 1200 subjects. In Figure 2, with 50% of subjects under-reporting events by 50% and 1/6 over-reporting events by 30%, the required sample size to maintain a power of 80% would grow to approximately 675 subjects per group or a total sample size of 1350 subjects. Although not shown in the Figures, multiple other factors would also inflate the estimated sample size. For example, we assumed that the bias in self reports was not differential between the experimental groups but we note that a differential bias in reporting would further inflate the estimated sample size. The current simulation assumed a negative binomial distribution in the count variables and did not consider the possibility of an excessive number of zero events. If the analysis considered the likely inflation of zero events in the distribution of events such as hospitalizations or nursing home stays, this would also inflate the estimated sample size.

While the simulation above focuses on the power and sample size implications of biased patient self-reports, over-reporting of care utilization also biases the estimation of treatment effect. In other words, patient over-reporting of care utilization affects not only the power of the test but also the estimation of the treatment effect. Figure 1 focuses on the power consequences of over and under-reporting. When a significant number of the study subjects over-report their utilization, the analytical power appears to be inflated (improved). But this power inflation comes at the expense of increased bias in treatment effect estimation. Thus, bias in self-reporting translates into bias in the estimated effect size.

Except for the no bias reporting scenario, all settings with under or over reporting also resulted in biased treatment effect estimates. In the scenario where 1/3 had no bias, 1/3 had under reporting and 1/3 had over reporting, the magnitude of the average bias was 4.5% of the true treatment effect. In the scenario where 1/3 had no bias, 1/2 had under reporting and 1/6 had over reporting, the magnitude of the average bias was 8% of the true treatment effect. In the scenario where 1/3 had no bias, 1/6 had under reporting and 1/2 had over reporting, the magnitude of the average bias was 17% of the true treatment effect. Notably, the scenario where 1/3 of patients had no reporting bias, 1/6 had under reporting and 1/2 had over reporting (which appeared to have the best power) actually had the worst level of bias in treatment effect estimate. In other words, the power inflation resulting from misreporting was achieved at the expense of substantially increased treatment effect bias. For Figure 2, the corresponding average treatment effect bias is 3%, 18%, and 13%.

DISCUSSION

Self-reports of health care use and hours spent in caregiving have been used to estimate the total cost of care for older adult with Alzheimer's disease.^{5,8,12,13,32} These reports consistently demonstrate higher costs of care for these patients relative to older adults without dementia. These self-reports also demonstrate the substantial cost of informal caregiving in the home as well as the cost of professional caregiving in the home or the

nursing home. In these prior studies, respondents typically report events such as hospitalizations or hours spent in caregiving over a fixed time interval such as 6 months or 1 year. These events are then translated into costs based on market estimates or other metrics relevant to the particular study and depending on the research team's decisions on assuming a societal perspective versus another economic perspective. Given the large differences in health care utilization between persons with dementia compared to those without dementia (including a large percentage of older adults with no hospitalizations or nursing home stays in a given year), this method of self-reported utilization provides valid estimates of the overall cost of dementia. However, it is less clear that self-reports of use are sufficiently accurate to allow comparisons between experimental groups of persons with Alzheimer's disease. Using data from an actual Alzheimer's disease clinical trial and a comprehensive database capturing health care utilization across the continuum of care, we demonstrate the magnitude of inaccuracy in self-reports. Caregivers tend to under-report use although errors in over-reporting also exist. We show that this bias in self-reports can substantially reduce statistical power in randomized trials and thereby increase the risk of a Type II error and these errors can also produce treatment effect bias that could produce inaccurate conclusions about the study results.

The use of self-reports rather than actual claims data is often justified through an argument that self-report data are less expensive to obtain. If the use of self-reports requires a larger sample size, however, self-report data may ultimately require a more expensive clinical trial. Most clinical trials to date have been powered based on outcomes such as cognition or behavior.^{16,17} These continuous variables, measured directly by the research team, often allow adequate statistical power at smaller sample sizes. Indeed, the typical Alzheimer's disease trial identified in recent systematic reviews had a sample size of less than 200 subjects.^{6,33–35} In additional or secondary analyses, researchers may use self-report data collected during these trials to ascertain whether the intervention reduced health care use or costs or delayed institutionalization. In studies enrolling fewer than 350 subjects per group, utilization counts would have to drop more than the 20% simulated in the current study in order to demonstrate a statistically significant difference between groups. In cost effectiveness analyses, error in self-report may be further magnified because events are translated into average costs per event. Cost data are highly skewed with large numbers of persons with zero or near zero costs. Thus, in the setting of small sample sizes, one subject over- or under-reporting estimated high-cost events can have a substantial impact on the study conclusions. In addition, differential bias in self-reports, particularly in the timing of events such as nursing home use, could result from respondents' perceptions of social desirability. A differential bias in self-reports could occur, for example, in a study where the intervention cannot be double-blinded such as an intervention designed to delay institutionalization through caregiver education and support.

In 1999, we reported on the limitations of using self-report data to investigate health services use among a community-based sample of older adults.²⁵ That study relied on a very different population that was a stratified random sample of 422 patients aged 60 and older who had visited the urban public health care system at least once in the previous 3 months. We completed a telephone to obtain subjects' self-reported health care use and compared these self-reports with data that were limited to local electronic medical records. Comparing

self-report data to the electronic record data, 24% of older adults with a hospitalization in the prior 12 months failed to report the episode and 28% of those with an emergency room visit failed to report the episode. Like the current study, self-reports of frequency of these services were substantially under-reported. In 2006, Bhandari and Wagner published a systematic review of 42 studies reporting on the accuracy of self-reported utilization data. The authors concluded that self-report data are of “variable accuracy” and that accuracy is affected by both telescoping and memory decay. Telescoping occurs when events are remembered to have occurred but respondents place the event in the wrong time period. The authors found that telescoping occurred regardless of the recall time frame and did not produce a consistent bias. The review also revealed that under-reporting was the most frequent problem and that under-reporting is made worse in the setting of increased health care use and older age. In the Bhandari and Wagner review, there were no studies that assessed the accuracy of proxy self-reports of health care use among older adults with dementia.

The term “under-reporting” can be viewed from different perspectives because the data are dominated by zero values and it is not necessary clear whether under-reporting one event in the context of one episode of care is more or less important than under-reporting one event in the context of many episodes of care. We have attempted to present and analyze the data in the manner most consistent with how the data are typically handled in a clinical trial and subsequent cost-effectiveness analyses. Extant studies rarely compare mean days of use for items like hospital, nursing home, or home health care because the distribution is severely skewed and dominated by zeros and this skewness is made more problematic by the relatively low number of enrolled subjects in a typical Alzheimer disease trial. The analyses in those studies, particularly for the nursing home outcomes which has been viewed as the most salient, is either “time to first use” or “any use”

The timing of nursing facility use has become an important metric for potential cost savings because delaying institutionalization by several months could amount to important cost savings. However, the results of our study suggest that proxy self-respondents may have difficulty in identifying an episode of nursing facility care and its precise timing or dates. Part of this difficulty may come from the vagaries inherent in terms such as “nursing home” and “institutionalization”. In the current environment in the United States, the same facility might provide for long-term care, subacute rehabilitation, hospice care, or respite care. We have previously documented the dynamic movement of patient with dementia from home to hospital to nursing home and back to home.¹¹ Over the time period of this study, subjects could also have spent time in “hospital-based long-term care”, “long-term acute care facilities”, assisted living settings, and rehabilitation centers. Because the term “nursing home” connotes both permanent institutionalization and short-term rehabilitation and because similar care can be provided in any of these settings and the home, respondents could understandably be confused by these terms. It is unlikely that survey questions completed by lay respondents could keep abreast of the changing landscape of nursing facilities and their capabilities.

This study has limitations. First, these results may not generalize to nationally representative samples of older adults with dementia. This is a small sample from an urban public health

care system but the sample size is similar to the majority of prior clinical trials studying older adults with dementia. Second, although we have assembled a comprehensive database of health care use across the care continuum, we would not necessarily have identified private insurer pay or self-pay ambulatory care visits or hospitalizations. The database does capture private insurer pay or self-pay nursing facility use or home health care use. Third, the database we assembled is expensive and the cost of collecting such claims data may be greater than the cost of expanding a study's sample size.

We conclude that proxy self-reports of health care use appear to contain significant error and the magnitude of the error is sufficient to reduce statistical power and introduce bias into treatment effect results. Gathering actual utilization data or adjudicating proxy respondent's self-reports should be encouraged for those trials where utilization is a primary outcome.

Acknowledgments

Supported by NIA grant R01 AG031222 and K24 AG024078

References

1. Schneider LS. Assessing outcomes in Alzheimer disease. *Alzheimer Dis Assoc Disord.* Aug; 2001 15(Suppl 1):S8–18. [PubMed: 11669509]
2. Courtney C, Farrell D, Gray R, et al. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet.* 2004; 363(9427):2105–2115. [PubMed: 15220031]
3. Wimo A. Cost effectiveness of cholinesterase inhibitors in the treatment of Alzheimer's disease: a review with methodological considerations. *Drugs Aging.* 2004; 21(5):279–295. [PubMed: 15040756]
4. Pinquart M, Sorensen S. Helping caregivers of persons with dementia: which interventions work and how large are their effects? *Int. Psychogeriatr.* Dec; 2006 18(4):577–595. [PubMed: 16686964]
5. Oremus M. Systematic review of economic evaluations of Alzheimer's disease medications. *Expert review of pharmacoeconomics & outcomes research.* Jun; 2008 8(3):273–289. [PubMed: 20528379]
6. Spijker A, Vernooij-Dassen M, Vasse E, et al. Effectiveness of nonpharmacological interventions in delaying the institutionalization of patients with dementia: a meta-analysis. *Journal of the American Geriatrics Society.* Jun; 2008 56(6):1116–1128. [PubMed: 18410323]
7. Bond M, Rogers G, Peters J, et al. The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease (review of Technology Appraisal No. 111): a systematic review and economic model. *Health Technol. Assess.* 2012; 16(21):1–470. [PubMed: 22541366]
8. Zhu CW, Sano M. Economic considerations in the management of Alzheimer's disease. *Clin Interv Aging.* 2006; 1(2):143–154. [PubMed: 18044111]
9. Weimer DL, Sager MA. Early identification and treatment of Alzheimer's disease: social and fiscal outcomes. *Alzheimers Dement.* May; 2009 5(3):215–226. [PubMed: 19362885]
10. Jones R, Sheehan B, Phillips P, et al. DOMINO-AD protocol: donepezil and memantine in moderate to severe Alzheimer's disease - a multicentre RCT. *Trials.* 2009; 10:57. [PubMed: 19630974]
11. Callahan CM, Arling G, Tu W, et al. Transitions in Care for Older Adults with and without Dementia. *Journal of the American Geriatrics Society.* May; 2012 60(5):813–820. [PubMed: 22587849]
12. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. *The New England journal of medicine.* Apr 4; 2013 368(14):1326–1334. [PubMed: 23550670]

13. Quentin W, Riedel-Heller SG, Luppá M, Rudolph A, König HH. Cost-of-illness studies of dementia: a systematic review focusing on stage dependency of costs. *Acta Psychiatr Scand.* Apr; 2010 121(4):243–259. [PubMed: 19694634]
14. Geldmacher DS, Provenzano G, McRae T, Mastey V, Ieni JR. Donepezil is associated with delayed nursing home placement in patients with Alzheimer's disease. *Journal of the American Geriatrics Society.* Jul; 2003 51(7):937–944. [PubMed: 12834513]
15. Mauskopf J, Mucha L. A review of the methods used to estimate the cost of Alzheimer's disease in the United States. *Am J Alzheimers Dis Other Dement.* Jun; 2011 26(4):298–309. [PubMed: 21561991]
16. Oremus M, Aguilar SC. A systematic review to assess the policy-making relevance of dementia cost-of-illness studies in the US and Canada. *Pharmacoeconomics.* Feb 1; 2011 29(2):141–156. [PubMed: 21090840]
17. Geldmacher DS. Cost-effectiveness of drug therapies for Alzheimer's disease: A brief review. *Neuropsychiatric disease and treatment.* Jun; 2008 4(3):549–555. [PubMed: 18830441]
18. Ferris SH, Aisen PS, Cummings J, et al. ADCS Prevention Instrument Project: overview and initial results. *Alzheimer Dis Assoc Disord.* Oct-Dec;2006 20(4 Suppl 3):S109–123. [PubMed: 17135805]
19. Sano M, Zhu CW, Whitehouse PJ, et al. ADCS Prevention Instrument Project: pharmacoeconomics: assessing health-related resource use among healthy elderly. *Alzheimer Dis Assoc Disord.* Oct-Dec;2006 20(4 Suppl 3):S191–202. [PubMed: 17135812]
20. Wimo A, Gustavsson A, Jonsson L, Winblad B, Hsu MA, Gannon B. Application of Resource Utilization in Dementia (RUD) instrument in a global setting. *Alzheimers Dement.* Nov 8.2012
21. Leicht H, Heinrich S, Heider D, et al. Net costs of dementia by disease stage. *Acta Psychiatr Scand.* Nov; 2011 124(5):384–395. [PubMed: 21838738]
22. Feldman H, Gauthier S, Hecker J, et al. Economic evaluation of donepezil in moderate to severe Alzheimer disease. *Neurology.* Aug 24; 2004 63(4):644–650. [PubMed: 15326236]
23. Zhu CW, Scarmeas N, Torgan R, et al. Clinical features associated with costs in early AD: baseline data from the Predictors Study. *Neurology.* 2006; 66(7):1021–1028. [PubMed: 16606913]
24. Roberts RO, Bergstralh EJ, Schmidt L, Jacobsen SJ. Comparison of self-reported and medical record health care utilization measures. *J Clin Epidemiol.* Sep; 1996 49(9):989–995. [PubMed: 8780606]
25. Wallihan DB, Stump TE, Callahan CM. Accuracy of self-reported health services use and patterns of care among urban older adults. *Med Care.* Jul; 1999 37(7):662–670. [PubMed: 10424637]
26. Ritter PL, Stewart AL, Kaymaz H, Sobel DS, Block DA, Lorig KR. Self-reports of health care utilization compared to provider records. *J Clin Epidemiol.* Feb; 2001 54(2):136–141. [PubMed: 11166528]
27. Raina P, Torrance-Rynard V, Wong M, Woodward C. Agreement between self-reported and routinely collected health-care utilization data among seniors. *Health Serv Res.* Jun; 2002 37(3):751–774. [PubMed: 12132604]
28. Bhandari A, Wagner T. Self-reported utilization of health care services: improving measurement and accuracy. *Med. Care Res. Rev.* Apr; 2006 63(2):217–235. [PubMed: 16595412]
29. Callahan CM, Boustani MA, Unverzagt FW, et al. Effectiveness of collaborative care for older adults with Alzheimer disease in primary care: a randomized controlled trial. *JAMA.* May 10; 2006 295(18):2148–2157. [PubMed: 16684985]
30. Ferris SH, Mackell JA, Mohs R, et al. A multicenter evaluation of new treatment efficacy instruments for Alzheimer's disease clinical trials: overview and general results. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord.* 1997; 11(Suppl 2):S1–12. [PubMed: 9236947]
31. McDonald CJ, Overhage JM, Barnes M, et al. The Indiana network for patient care: a working local health information infrastructure. An example of a working infrastructure collaboration that links data from five health systems and hundreds of millions of entries. *Health Aff (Millwood).* Sep-Oct;2005 24(5):1214–1220. [PubMed: 16162565]

32. Zhu CW, Scarmeas N, Torgan R, et al. Clinical characteristics and longitudinal changes of informal costs of Alzheimer's disease in the community. *Journal of the American Geriatrics Society*. 2006; 54(10):1596–1602. [PubMed: 17038080]
33. Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane database of systematic reviews (Online)*. 2006; (1):CD005593.
34. Raina P, Santaguida P, Ismail A, et al. Effectiveness of cholinesterase inhibitors and memantine for treating dementia: evidence review for a clinical practice guideline. *Annals of internal medicine*. Mar 4; 2008 148(5):379–397. [PubMed: 18316756]
35. Kurz AF, Leucht S, Lautenschlager NT. The clinical significance of cognition-focused interventions for cognitively impaired older adults: a systematic review of randomized controlled trials. *Int. Psychogeriatr*. Jul 11.2011 :1–12. [PubMed: 20550746]

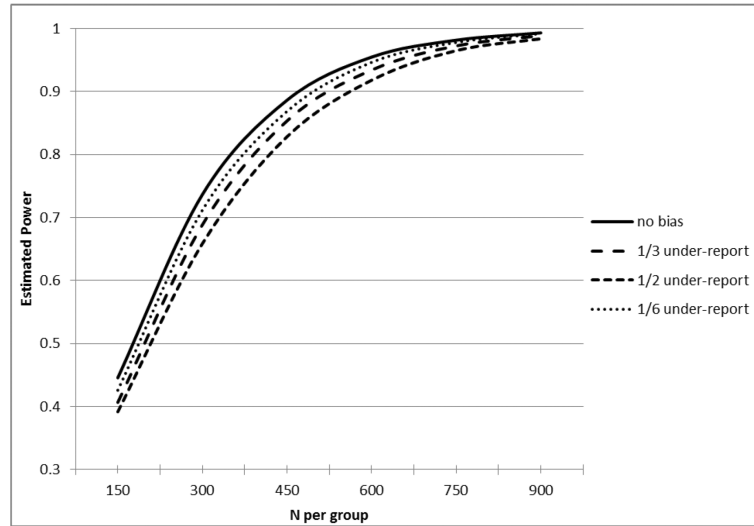


Figure 1. Estimated Power for an Equal-sized Two Group Comparison in Poisson Regression of a Count Outcome (assuming 20% reduction with $\log_e(.7)$ under-reporting effect).

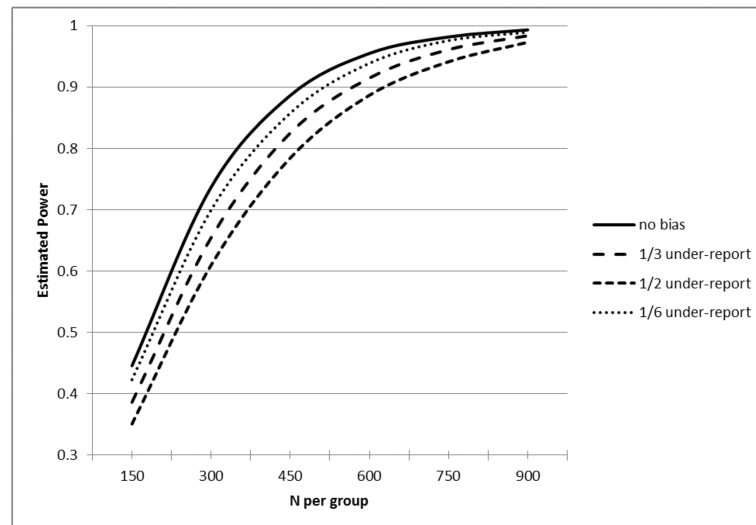


Figure 2. Estimated Power for an Equal-sized Two Group Comparison in Poisson Regression of a Count Outcome (assuming 20% reduction with $\log_e(.5)$ under-reporting effect).

Table 1

Baseline characteristics of study sample (N=100)

Variable	n (%) or mean (SD)
Patient characteristics	
Female, n (%)	66 (66.0)
Black, n (%)	39 (39.0)
Married, n (%)	23 (23.0)
Medicaid recipient, n (%)	36 (39.6)
Age, mean (SD)	76.6 (6.3)
Annual income, mean (SD), \$	9,735.3 (6,965.3)
Education, mean (SD), years	7.8 (3.6)
MMSE score, mean (SD)	17.8 (5.3)
No of medications taking, mean (SD)	5.2 (2.7)
Chronic disease score, mean (SD)	6.8 (3.5)
Caregiver characteristics	
Age, mean (SD)	54.4 (13.5)
Female, n (%)	85 (85.0)
Live with patient, n (%)	55 (55.0)
Relationship of caregiver to patient, n (%)	
Spouse	17 (17.0)
Child	54 (54.0)
Other	29 (29.0)

Table 2

Agreement between actual use and self-reports based on any use vs. no use—cumulative use overtime, N=100

Site of care	-6 -0 months	-6-6 months	-6-12 months	-6-18 months
Emergency Department				
Any use in claims, n (%)	39 (39.0)	47 (47.0)	56 (56.0)	63 (63.0)
Respondents reporting some use when none, n (%) [*]	17 (27.9)	19 (35.8)	19 (43.2)	19 (51.3)
Respondents reporting no use when some, n (%) [†]	22 (56.4)	22 (46.8)	20 (35.7)	19 (30.2)
Agreement between claims and self-report [‡]	0.16	0.17	0.21	0.18
Hospital				
Any use in claims, n (%)	14 (14.0)	21 (21.0)	25 (25.0)	29 (29.0)
Respondents reporting some use when none, n (%) [*]	10 (11.6)	12 (15.2)	16 (21.3)	14 (19.7)
Respondents reporting no use when some, n (%) [†]	3 (21.4)	3 (14.3)	3 (12.0)	2 (06.9)
Agreement between claims and self-report [‡]	0.55	0.61	0.57	0.65
Home Health Care				
Any use in claims, n (%)	07 (7.0)	13 (13.0)	17 (17.0)	21 (21.0)
Respondents reporting some use when none, n (%) [*]	9 (09.7)	8 (09.2)	11 (13.2)	12 (15.2)
Respondents reporting no use when some, n (%) [†]	4 (57.1)	2 (15.4)	3 (17.6)	2 (09.5)
Agreement between claims and self-report [‡]	0.25	0.63	0.58	0.64
Nursing Home				
Any use in claims, n (%)	02 (2.0)	03 (3.0)	03 (3.0)	06 (6.0)
Respondents reporting some use when none, n (%) [*]	6 (06.1)	8 (08.2)	11 (11.3)	8 (08.5)
Respondents reporting no use when some, n (%) [†]	1 (50.0)	2 (66.6)	2 (66.6)	2 (33.3)
Agreement between claims and self-report [‡]	0.20	0.13	0.09	0.40

* percentage refers to the number of persons reporting some use divided by the number with no actual use;

† percentage refers to the number of persons reporting no use divided by the number with some actual use;

‡ Kappa statistic

Table 3

Agreement between actual use and self-reports based on any use vs. no use—individual 6 month periods

Site of care	-6 -0 months	0-6 months	6-12 months	12-18 months
N	100	89	80	68
Emergency Department				
Any use in claims, %	39 (39.0)	30 (33.7)	34 (42.5)	26 (38.2)
Respondents reporting some use when none, % [*]	17 (27.9)	13 (22.0)	9 (19.6)	11 (26.2)
Respondents reporting no use when some, % [†]	22 (56.4)	16 (53.3)	15 (44.1)	14 (53.9)
Agreement between claims and self-report [‡]	0.16	0.25	0.37	0.20
Hospital				
Any use in claims, %	14 (14.0)	10 (11.2)	8 (10)	8 (11.8)
Respondents reporting some use when none, % [*]	10 (11.6)	8 (10.1)	9 (12.5)	5 (8.3)
Respondents reporting no use when some, % [†]	3 (21.4)	5 (50.0)	3 (37.5)	1 (12.5)
Agreement between claims and self-report [‡]	0.55	0.35	0.38	0.65
Home Health Care				
Any use in claims, %	07 (7.0)	10 (11.2)	9 (11.3)	9 (13.2)
Respondents reporting some use when none, % [*]	9 (09.7)	5 (06.3)	5 (07.0)	4 (06.8)
Respondents reporting no use when some, % [†]	4 (57.1)	4 (40.0)	4 (44.4)	1 (11.1)
Agreement between claims and self-report [‡]	0.25	0.51	0.46	0.72
Nursing Home				
Any use in claims, %	02 (2.0)	01 (1.1)	00 (0.0)	3 (04.4)
Respondents reporting some use when none, % [*]	6 (06.1)	2 (02.3)	3 (03.8)	0 (00.0)
Respondents reporting no use when some, % [†]	1 (50.0)	1 (100.0)	na	0 (00.0)
Agreement between claims and self-report [‡]	0.20	-0.02	na	1.00

na=not applicable

^{*} percentage refers to the number of persons reporting some use divided by the number with no actual use; for example, reporting of ED visits when none -6-0 months $17/(100-39)=27.9\%$

[†] percentage refers to the number of persons reporting no use divided by the number with some actual use; for example, reporting no ED visits when some -6-0 months $22/39=27.9\%$

[‡] Kappa statistic

Table 4

Agreement between actual use and self-reports based on volume of use

Site of care	6 mo period – -6 to 0 months	12 mo period – -6 to 6 months
Emergency Department		
Correctly report zero use, n (%)	44 (44.0)	34 (34.0)
Correctly report some use, n (%)	8 (8.0)	9 (9.0)
Under-report volume of use, n (%)	29 (29.0)	31 (31.0)
Over-report volume of use, (%)	19 (19.0)	26 (26.0)
Agreement between claims and self-report *	0.14	0.14
Hospital		
Correctly report zero use, n (%)	76 (76.0)	67 (67.0)
Correctly report some use, n (%)	7 (7.0)	12 (12.0)
Under-report volume of use, n (%)	3 (3.0)	5 (5.0)
Over-report volume of use, n (%)	14 (14.0)	16 (16.0)
Agreement between claims and self-report *	0.44	0.49
Home Health Care		
Correctly report zero use, n (%)	84 (84.0)	79 (79.0)
Correctly report some use, n (%)	2 (2.0)	7 (7.0)
Under-report volume of use, n (%)	5 (5.0)	4 (4.0)
Over-report volume of use, n (%)	9 (9.0)	10 (10.0)
Agreement between claims and self-report *	0.22	0.51
Nursing Home		
Correctly report zero use, n (%)	92 (92.0)	89 (89.0)
Correctly report some use, n (%)	1 (1.0)	1 (1.0)
Under-report volume of use, n (%)	1 (1.0)	2 (2.0)
Over-report volume of use, n (%)	6 (6.0)	8 (8.0)
Agreement between claims and self-report *	0.20	0.14

* Kappa statistic