August 25, 2020

Using Medicare claims in identifying Alzheimer's disease and related dementias

Siddharth Jain, DrPH[1,2]
Paul R. Rosenbaum, PhD[2,3]
Joseph G. Reiter, MS[1]
Geoffrey Hoffman, PhD[4,5]
Dylan S. Small, PhD[2,3]
JinKyung Ha, PhD[6]
Alexander S. Hill, BS[1]
David A. Wolk, MD[7]
Timothy Gaulton, MD[2,8,9]
Mark D. Neuman, MD[2,8,9]
Roderic G. Eckenhoff, MD[9]
Lee A. Fleisher, MD[2,8,9]
Jeffrey H. Silber, MD, PhD[1,2,9,10,11]

Affiliations:
[1]Center for Outcomes Research, Children's Hospital of Philadelphia, Philadelphia, PA
[2]The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA
[3]Department of Statistics, The Wharton School, The University of Pennsylvania, Philadelphia, PA
[4]Department of Systems, Populations and Leadership, University of Michigan School of Nursing, Ann Arbor, MI, USA;
[5]University of Michigan's Institute for Healthcare Policy and Innovation, Ann Arbor, MI, USA
[6]Division of Geriatrics/Institute of Gerontology, University of Michigan, Ann Arbor, MI, USA
[7]Department of Neurology, The Perelman School of Medicine, The University of Pennsylvania
[8]Center for Perioperative Outcomes Research and Transformation, The University of Pennsylvania, Philadelphia, PA
[9]Department of Anesthesiology and Critical Care, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[10]The Departments of Pediatrics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[11]Department of Health Care Management, The Wharton School, The University of Pennsylvania, Philadelphia, PA

Address correspondence: Dr. Siddharth Jain, Center for Outcomes Research, Children's Hospital of Philadelphia, Roberts Building, 2716 South Street, Room 5140, Philadelphia, PA 19146-2305. (T) 267-425-1712; (F) 215-590-2378; Email: jains@email.chop.edu.

Declaration of Interest: None

Short title: Claims Derived Alzheimer's Identification

ABSTRACT: (141/150)

INTRODUCTION: This study develops a measure of Alzheimer's disease and Related Dementias (ADRD) using Medicare claims.

METHODS: Validation resembles the approach of the American Psychological Association, including (i) content validity, (ii) construct validity, and (iii) predictive validity.

RESULTS: We found that four items -- a Medicare claim recording ADRD 1-year ago, 2-years ago, 3-years ago, and a total stay of 6-months in a nursing home -- exhibit a pattern of association consistent with a single underlying ADRD construct, and presence of any two of these four items predict a direct measure of cognitive function and also future claims for ADRD.

DISCUSSION: Our four items are internally consistent with the measurement of a single quantity. The presence of any two items do a better job than a single claim when predicting both a direct measure of cognitive function and future ADRD claims.

Keywords: (5 to 15 keywords) Alzheimer's, Dementia, ADRD, Cognitive impairment, Medicare, Medicare Claims, Administrative data, Health and Retirement Study

1. INTRODUCTION (3,538/3,500)

There will be around 5.8 million cases of Alzheimer's Disease by 2020, with the number expected to double by 2040 in the USA [1]. Research [2-5] examining Alzheimer's disease and related dementias (ADRD) often relies on diagnoses from Medicare administrative claims.

This study develops a new claims-based measure of ADRD that is readily computed for millions of people covered by Medicare. Our measure is checked for validity in several ways: (i) content validity: use of appropriate ICD-9/10 codes, (ii) construct validity: components of the measure exhibit internal patterns of association consistent with several

indicators of a single underlying disorder, (iii) other claims-based empirical checks: reliable anticipation of future Medicare claims for ADRD, and consistency with nursing home assessments of cognitive impairment, and finally, (iv) predictive validity: the measure is appropriately correlated with a direct, external measure of cognitive function, not based on Medicare claims, assessed using the Health and Retirement Study (HRS).

2. METHODS

2.1. Conceptual Model

We follow the standard conceptual model for validation of a psychological measure, namely content, construct, and predictive validity [6-8] and we do these employing concepts from nonparametric item response theory [9-16]. Many articles define ADRD using Medicare claims with the presumption that one [3-5,17-19] ADRD claim or two [20,21] claims close together in time suffices to establish the diagnosis of ADRD. In contrast, in this study, we examine a definition using two claims of ADRD widely separated in time, and examine whether this definition exhibits greater external validity in predicting a direct measure of cognitive impairment, a greater ability to predict future claim of ADRD, and internal consistency reflective of a single underlying disorder.

Dementia is not a transient cognitive problem; so transient evidence of a cognitive problem is insufficient to justify categorizing a patient as demented. While a hospital admission may produce a bundle of ADRD diagnoses codes, such bundled codes may reflect a transient cognitive problem, perhaps all derived from a single evaluation. We will demonstrate that multiple codes in a short time interval of time are too correlated with one another—too idiosyncratic—to be compatible with a single construct producing ADRD codes over several years [22]. Our definition requires that a claim of ADRD be confirmed by another process, either a second claim widely separated in time or a lengthy stay in a nursing

home. Close to two-thirds of Medicare beneficiaries living in nursing homes have ADRD [23].

2.2 Patient Population and Available Data

To develop our claims-based algorithm, we used Medicare Inpatient, Outpatient, Carrier, Skilled Nursing Facility (SNF), Home Health Agency (HHA), and Long-Term Care Minimum Data set (MDS) files of all Medicare beneficiaries from years 1999-2016. The data to develop the algorithm consisted of 697,870 patients, which is a 1% random sample of all Medicare beneficiaries aged 65 years and older. Beneficiaries were excluded if, in their look-back period of three years, they were not fully enrolled in Part A and Part B; were enrolled in Medicare Advantage for at least a month (Medicare Advantage enrollees were excluded because Medicare does not make claims data available for these patients); or did not have any inpatient, outpatient, or physician claims. After our definition was developed, we applied it to an external data set that including Medicare patients in the 2002-2014 Health and Retirement Study (HRS) surveys. Our validation sample included 4,291 patients aged 75; 3,431 patients aged 80; and 2,489 patients aged 85.

2.3. Study Design Overview

2.3.1. Step 1: Content Validity: Selecting a list of indicators of ADRD: In the current context, content validity means using correct diagnostic codes and recognizing incorrect codes.

To ensure the correct list of codes, we began with a list of ADRD diagnostic codes (Table 1) used in the literature (Appendix A). The Chronic Condition Warehouse (CCW) has defined ADRD [18] with the presence of at least one claim with the diagnosis codes in the CCW list (Appendix A) in Inpatient, Outpatient, Carrier, SNF, and HHA files in three years [5]. Several studies have used CCW definition of ADRD [24-27]. We also considered codes for

other medical conditions, such as other degenerative conditions, delirium (Appendix B), and 6-months nursing home stay (a total number of 180 days from all the stays in a nursing home) from the MDS file [28] in the three years look-back period.

2.3.2. Step 2: Construct Validity: Developing a coding algorithm to define ADRD based on selected indicators: Construct validation is supported by a check that the components of the measure exhibit a pattern of associations that measure a single underlying disorder. Since ADRD is not a transient condition, we checked if codes dispersed through multiple years are a better marker of ADRD than an equal number of codes in a single year. Specifically, we identify a patient as having ADRD as of date *d* if during the three years look-back either: (i) there was at least one claim with a code for ADRD in at least two different years, or (ii) there was at least one claim with a code for ADRD plus there was a total stay of 6-months in a nursing home. Taylor et al. suggest that three years of data are favorable [5].

Our measure of ADRD is intended to distinguish, as much as possible, ADRD from other diagnostic codes reflective of other diseases. We used item response theory to test whether our four indicators exhibit the pattern of associations consistent with multiple indicators of one underlying disorder—that is, consistent with unidimensionality. This pattern says that every pair of two indicators exhibits a nonnegative association given any function of the remaining indicators [10,12]. Three of our four indicators were the presence of at least one ADRD code (Table 1) i) 1-year ago (0-365 days), ii) 2-years ago (366-730 days), iii) 3-years ago (731-1095 days), and our fourth indicator was iv) a total stay of 6-months in a nursing home. We studied these relationships at three ages, 75, 80, and 85 years, looking back three years from a patient's birthday.

Also, we constructed three similar indicators for other degenerative conditions and delirium codes (Appendix B) to examine whether Medicare claims can distinguish ADRD from other cognitive disorders.

If various items of information, such as Medicare claims mentioning specific ICD-9/10 codes, are indicative of a single disorder such as the degree of dementia, then those items should exhibit certain patterns of association or interdependence [10,12]. We used claims data to check for this pattern, rejecting initial versions of the measure that violated the pattern. These checks led to the conclusion that our 4th indicator, a total stay of 6-months in a nursing home, acts as a check on information from claims: it exhibits the appropriate pattern of associations when conjoined with annual indicators of at least one claim mentioning ADRD. Importantly, ADRD is never diagnosed based on the 4th indicator alone; there must also be a claim for ADRD.

2.3.3. Description of the claims-based coding algorithm to define ADRD

Our ADRD definition is based on 4 binary indicators. A patient gets one point for the presence of each indicator. The score is the sum of these points. A confirmed diagnosis of ADRD is a score of ≥2.

2.3.4. Further Empirical Checks: The Cognitive Performance Scale (CPS) [29] is from the MDS 2.0 files and is an assessment of cognitive function among nursing home patients [30]. The CPS score ranges from 0 (intact cognition) to 6 (very severe impairment) [29]. We checked our binary ADRD indicator measure against the CPS on the subset of patients in nursing homes. The CPS is available only in nursing homes, and our ADRD indicator takes account of whether a person has spent a total of 6-months in a nursing home, so this check is interesting but imperfect.

2.3.5. Step 3: Predictive Validity: External validation of the ADRD coding algorithm based on HRS-Medicare data: The HRS is an ongoing nationally representative survey that collects biennial data on subjects' cognition and functional status [31]. We validated our ADRD indicator with direct and proxy assessments of cognition in the HRS. Cognitive functioning in

the HRS was assessed using an adapted version of the Telephone Interview for Cognitive Status (TICS) [32]. The cognition measure for self-respondents was scored on a scale of 0-35 [33]. A proxy respondent can provide answers about symptoms of the survey participant who cannot take survey due to functional and cognitive limitations. Langa-Weir [34] have used alternative assessments (from proxy respondents and interviewers) to assess cognition.

2.4. Statistical Methods

2.4.1. Using multiple years of codes for the coding algorithm: To test if codes dispersed through multiple years are a better marker of ADRD than single-year code, we constructed Cox proportional hazard models [35]. Patients were classified as having: (i) no ADRD codes, (ii) ADRD codes in a single year, (iii) ADRD codes in multiple (≥2) years, in the three-year lookback. The model examined the time to another ADRD code in the future after the end of the lookback period, adjusting for the total number of ADRD codes because some patients accumulate many more codes than other patients. Does temporal dispersion of codes add information beyond their number?

In yet another check, we used simple regression models regressing CPS scores on the temporal dispersion of the ADRD codes in the three-year lookback period. We estimated the time differences between a patient's birthday and the dates on which the patient received ADRD diagnostic codes. The temporal dispersion of codes was defined as the standard deviation of the length of these time periods. The standard deviation reflects whether the codes are dated close together in time or dispersed over several years.

2.4.2. Methods to test construct validity: Our measure of ADRD is the presence of two of our four binary items or indicators. The four binary events are recorded in a $2^4 = 16$-fold table of counts. For construct validation, we determined whether that 16-fold table exhibits the pattern of associations consistent with four measures of a single underlying disorder [10,12]. A

single disorder implies a nonnegative association between any pair of two items conditionally given the total score on the two remaining items—a *2x2x3* table derived from the 16-fold table—and nonparametric item response theory checks this using the Mantel-Haenszel statistic applied to the *2x2x3* table [10,12]. Following Campbell's notion of discriminant validation [7], we did these steps also incorporating indicators of other degenerative conditions and delirium—now unidimensionality is lost as negative partial associations are produced indicative of more than one underlying disorder.

Technically, the term "single disorder" refers to a mathematical structure called "monotone latent unidimensionality," which contends that our four binary indicators agree with each other because they are each fallible measures of one underlying quantity, presumably the degree of dementia. Monotonicity means that more indicators suggest a greater degree of dementia. If this mathematical structure were true, it would have various consequences that can be checked in observable data, and our analyses of the 2^4 = 16-fold table of counts are checks of these consequences.

Additionally, we checked that each of the four indicators increases the probability of ADRD given the total of the other three indicators, a pattern called monotonicity [16].

2.4.3. Methods for empirical checks: For the checks with the CPS scores from nursing home assessments, we used the Goodman-Kruskal gamma (γ) coefficient [36], which can measure the association between the ordinal categories based on CPS scores and the binary categories based on claims-based ADRD definition. We also calculated the Probability of Concordance (C) from γ as $C = (1 + \gamma)/2$ [36] which is the probability that the two measures order two patients in the same way.

2.4.4 Methods for external validation: The cognitive impairment threshold for self-respondents was a score of 8 or less as suggested by Herzog and Wallace in their initial analyses of HRS [37] and has been used in the literature [3,38-40], and a score of 6 or more for assessment of cognition by proxies and interviewers, where a higher score is classified as cognitive impairment [34]. We compared the frequency of patients who were identified as cognitively impaired by either HRS self-response (score ≤8) or proxy/interviewer reporting (score ≥6) to the frequency of patients defined as ADRD based on a ≥2 score from our claims-based algorithm.

3. RESULTS

3.1: Multiple years of codes: Cox-proportional hazard models (Table 2) examined whether the temporal dispersion of ADRD codes predicts future ADRD codes. We found that ADRD codes dispersed over a three-year lookback were more predictive of future ADRD codes than were an equal number of ADRD codes occurring at almost the same time. At age 75, patients with codes in multiple years as opposed to a single year have a 4-times higher hazard (unadjusted HR=4.386, p-value <0.0001; adjusted HR=2.297, p-value <0.0001) of getting a diagnostic code later in life. Models for other ages were similar.

Regression models found greater temporal dispersion of ADRD codes is significantly associated with a greater cognitive decline on CPS scores (p-value <0.001) (Appendix D). An analysis using the stratified Wilcoxon rank-sum test supported the same conclusion (p-value <0.001) (Appendix E).

3.2: Construct validity: Do our four items measure a single disorder? As seen in Table 3, there is a nonnegative association between any two items given the total score on the remaining two items, a pattern consistent with four items measuring a single disorder [10,12]. All the odds ratios between all items pairs were significantly greater than one (p-value

<0.0001) for all age groups. In sharp contrast, adding other degenerative conditions codes creates a second dimension, with some conditional odds ratios significantly below one for age 75 (p-value <0.05), and age 80 and 85 (<0.0001) (appendix F). Similarly, adding delirium codes yields a second dimension evident from odds ratios significantly less than one for age 75 (p-value <0.05), and age 80 and 85 (<0.0001) (appendix G).

The four items exhibit the expected increasing relationship between four fallible measures of one underlying disorder: they exhibit monotonicity [16]. For instance, at age 75, the probability of a patient having the item ADRD code 1-year ago is 1.82% if none of the other three items were present in the look-back, but this rises to 90.22% if all of the other three items were present—moreover, probability of concordance between this item and the score on the other three is estimated to be 97.7% (appendix H). The four items exhibit an appropriately high level of internal consistency. Results were similar at other ages.

3.3: Internal checks of claims data with other unused claims data: Inside nursing homes, patients flagged by claims as ADRD (score ≥2) were more likely to have a higher CPS score than those who were not flagged (score ≤1) (Appendix I). Inside nursing homes, the probability of concordance between these two measures was 87.21%, 85.67%, and 85.33% for age groups 75, 80, and 85 respectively.

3.4: Predictive validity: The distribution of HRS scores is shown in Table 4. The probability of concordance between HRS Self-respondents score (lower score indicates poor cognition) and our algorithm score was 83.80%, 81.06%, and 81.04%, and for proxy respondents (higher score indicates poor cognition) it was 89.24%, 83.43 % and 82.63% for age 75, 80, and 85, respectively.

Comparing our ADRD definition with cognitive impairment defined using HRS responses, we found that among 4,291 HRS respondents age 75 years, the odds of having

ADRD by our algorithm (score ≥2) was 49.01 times higher (95% CI 33.81, 71.04) for those who were defined as cognitively impaired compared to those who were not cognitively impaired in HRS. The odds ratios were 35.49 (95% CI 26.00, 48.45) for respondents aged 80 years and 20.81 (95% CI 15.58, 27.81) for respondents aged 85 years (Table 5). The odds ratios in Table 5 reflecting agreement between HRS and our algorithm are higher at every age than a definition that uses a single claim (Appendix J).

Validity refers to the ability of a proposed measure to predict some other external criterion measure. Here, the external measure is from the HRS. Criterion-related construct validity means demonstrating that each component of a proposed measure makes a nonnegative contribution to prediction—essentially that no component of our four-component measure is detracting from prediction by measuring something else. We checked for criterion-related construct validity [11], meaning that each of our four indicators makes a positive contribution to predicting the HRS given the total on the other three indicators. Each item contributes something unique, specifically, the Mantel-Hanszel test shows that there is a positive and significant (p-value <0.001) association between cognitive impairment based on HRS surveys and each of four items given the total score on the remaining three items (Appendix K).

We extended our primary list of codes (Table 1), with additional diagnostic codes (Appendix C) that have been used in various studies [19-21,24,41-45] to validate our measure of ADRD. We found similar results in the concordance between HRS scores and the scores from our algorithm (Appendix L), and the odds ratios obtained from our algorithm vs. the HRS definition of cognitive impairment (Appendix M).

Dynamic method: The evaluation has focused on ADRD at three ages, 75, 80 and 85; however, it will often be of interest to date the first symptoms of ADRD. In an alternate dynamic method, we computed our 4-item claims score on each day when a patient received

an ADRD code. Each date was given a score 0-4 based on the presence of our four indicators in a window of three-year look back and a three-year look forward period. The patient was defined as having ADRD on the date of the first code that led to score ≥2.

Comparing these dynamic scores to the HRS, we found that the probability of concordance between HRS Self-respondents score and our algorithm score was 85.06%, 82.75%, and 81.59%, and for the HRS proxy respondents' scores the probability of concordance was 89.38%, 85.26% and 82.24% for age 75, 80, and 85, respectively (Appendix N). We also found that the odds of having ADRD was 35.62 (25.21, 50.33), 34.34 (25.41, 46.41), and 20.19 (15.14, 26.93) times higher for age 75, 80 and 85, respectively, in those who were defined as cognitively impaired compared to those who were not in HRS (Appendix O).

Again, we used our extended list of diagnostic codes to compute scores based on the dynamic method and found similar results (Appendix P and Q).

4. DISCUSSION

This study identifies ADRD patients using Medicare claims. We find that several ADRD codes dispersed over a long period of time aid diagnosis of ADRD, as does a lengthy total amount of time spent in a nursing home. We also find that patients who have ADRD codes spread through multiple years in the lookback are more likely to receive diagnosis codes in their future claims compared to patients who have codes only in a single year after controlling for the total number of codes. The claims-based algorithm proposed in this study uses a definition of the presence of at least two of the four indicators (a score of ≥2) spread out in a period of three years to be identified as ADRD.

Our four indicators appear to be unidimensional and are consistent with each other in measuring one underlying disorder; moreover, each indicator aids prediction of a direct measure of cognitive decline beyond what the other three can do. Internal validation of the

indicators with the cognitive assessment in nursing homes also showed that patients with ADRD (≥2 scores) based on the algorithm have higher CPS scores compared to those who do not have ADRD.

The claims-based coding algorithm was externally validated with HRS cognitive scores. Patients defined as ADRD on the algorithm had worse cognitive scores in the HRS. Also, the odds of having ADRD by our algorithm were higher among those defined as cognitively impaired in the HRS survey.

Previous work that compared assessment of ADRD in surveys with evidence of ADRD in Medicare claims found poor or fair agreement between the two [3,21,46,47]. However, these studies did not account for the dispersion of codes through multiple years and used the presence of a single claim of ADRD as a diagnosis in claims. Another study using ADAMS (Aging Demographics and Memory Study) data, a cohort of HRS, also used a single claim of ADRD: it found that the use of a single Medicare claim results in an over-count of the true prevalence [17]. One ADRD claim is not sufficient: it might reflect transient confusion rather than dementia. Several claims filed at nearly the same time may reflect one assessment recorded several times. Claims that only appear close to each other in time suggest a transient condition, and perhaps the automated activities of a single person entering multiple records into the Medicare system. ADRD is a long-lasting condition and claims from patients with ADRD spread over a significant period of time strengthen the diagnosis of ADRD.

There is no gold standard for measuring ADRD that is applicable in studies of large numbers of patients. Nonetheless, our four-item, claims-based scale exhibits appropriate internal and external consistency, and a score of 2 or more predicts both future ADRD claims and direct indicators of cognitive function from the HRS. Moreover, our indicator of ADRD can be calculated at a negligible cost for millions of Medicare patients.

ACKNOWLEDGEMENTS

This research was supported by the National Institute on Aging (grant number RF1-AG-055390).

REFERENCES

[1] Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology. 2013;80:1778-83.

[2] Rocca WA, Petersen RC, Knopman DS, Hebert LE, Evans DA, Hall KS, et al. Trends in the incidence and prevalence of Alzheimer's disease, dementia, and cognitive impairment in the United States. Alzheimers Dement. 2011;7:80-93.

[3] Ostbye T, Taylor DH, Jr., Clipp EC, Scoyoc LV, Plassman BL. Identification of dementia: agreement among national survey data, medicare claims, and death certificates. Health Serv Res. 2008;43:313-26.

[4] Akushevich I, Yashkin AP, Kravchenko J, Ukraintseva S, Stallard E, Yashin AI. Time trends in the prevalence of neurocognitive disorders and cognitive impairment in the United States: The effects of disease severity and improved ascertainment. J Alzheimers Dis. 2018;64:137-48.

[5] Taylor DHJ, Fillenbaum G, G., Ezell ME. The accuracy of Medicare claims data in identifying Alzheimer's disease. J Clin Epidemiol. 2002;55:929-37.

[6] Messick S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. Am Psycol. 1995;50:741-49.

[7] Campbell DT. Recommendations for APA test standards regarding construct, trait, or discriminant validity. Am Psycol. 1960;15:546-53.

[8] Sechrest L. Validity of measures is no simple matter. Health Serv Res. 2005;40:1584-604.

[9] Sijtsma K, Molenaar IW. Introduction to Nonparametric Item Response Theory. Jaeger RM, ed. Vol. 5. Thousand Oaks, CA: Sage Publications, Inc.; 2002.

[10] Holland PW, Rosenbaum PR. Conditional association and unidimensionality in monotone latent variable models. Ann Stat. 1986;14:1523-43.

[11] Rosenbaum PR. Criterion-related construct validity. Psychometrika. 1989;54:625-33.

[12] Rosenbaum PR. Testing the conditional independence and monotonicity assumptions of item response theory. Psychometrika. 1984;49:425-35.

[13] Stout W. Nonparametric item response theory: A maturing and applicable measurement modeling approach. Appl Psychol Meas. 2001;25:300-06.

[14] Molenaar IW. Thirty years of nonparametric item response theory. Appl Psychol Meas. 2001;25:295-99.

[15] Junker BW, Sijtsma K. Nonparametric item response theory in action: An overview of the special issue. Appl Psychol Meas. 2001;25:211-20.

[16] Junker BW, Sijtsma K. Latent and manifest monotonicity in item response models. Appl Psychol Meas. 2000;24:65-81.

[17] Taylor DH, Jr., Ostbye T, Langa KM, Weir D, Plassman BL. The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. J Alzheimers Dis. 2009;17:807-15.

[18] Chronic Conditions Data Warehouse. Condition Categories. 2017. https://www.ccwdata.org/web/guest/condition-categories. Accessed November 29, 2017.

[19] Lee E, Gatz M, Tseng C, Schneider LS, Pawluczyk S, Wu AH, et al. Evaluation of Medicare claims data as a tool to identify dementia. J Alzheimers Dis. 2019;67:769-78.

[20] Desai U, Kirson NY, Ye W, Mehta NR, Wen J, Andrews JS. Trends in health service use and potentially avoidable hospitalizations before Alzheimer's disease diagnosis: A matched, retrospective study of US Medicare beneficiaries. Alzheimers Dement (Amst). 2019;11:125-35.

[21] Chen Y, Tysinger B, Crimmins E, Zissimopoulos JM. Analysis of dementia in the US population using Medicare claims: Insights from linked survey and administrative claims data. Alzheimers Dement (N Y). 2019;5:197-207.

[22] Rosenbaum PR. Item bundles. Psychometrika. 1988;53:349-59.

[23] Gaugler JE, Yu F, Davila HW, Shippee T. Alzheimer's disease and nursing homes. Health Aff (Millwood). 2014;33:650-7.

[24] White L, Fishman P, Basu A, Crane PK, Larson EB, Coe NB. Medicare expenditures attributable to dementia. Health Serv Res. 2019;54:773-81.

[25] Zhu CW, Ornstein KA, Cosentino S, Gu Y, Andrews H, Stern Y. Misidentification of dementia in Medicare claims and related costs. J Am Geriatr Soc. 2019;67:269-76.

[26] Matthews KA, Xu W, Gaglioti AH, Holt JB, Croft JB, Mack D, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged >/=65 years. Alzheimers Dement. 2019;15:17-24.

[27] Dutcher SK, Rattinger GB, Langenberg P, Chhabra PT, Liu X, Rosenberg PB, et al. Effect of medications on physical function and cognition in nursing home residents with dementia. J Am Geriatr Soc. 2014;62:1046-55.

[28] Research Data Assistance Center. Long term care minimum data sets (MDS) - Swing-bed 2.0. https://www.resdac.org/cms-data/files/mds-sb-2.0. Accessed July 26, 2019.

[29] Morris JN, Fries BE, Mehr DR, Hawes C, Phillips C, Mor V, et al. MDS Cognitive Performance Scale. J Gerontol. 1994;49:M174-82.

[30] Center for Medicare & Medicaid Services. Nursing Home Quality Initiative MDS 2.0. Last Update Date: 03/06/2015. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/NHQIMDS20.html. Accessed July 8, 2019.

[31] Institute for Social Research, University of Michigan. Health and Retirement Study. Data Products. 2019. https://hrs.isr.umich.edu/data-products. Accessed July 8, 2019.

[32] Brandt J, Spencer M, Folstein MF. The telephone interview for cognitive status. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. 1988;7:111-17.

[33] Ofstedal MB, Fisher GG, Herzog AR. Documentation of cognitive functioning measures in the health and retirement study. HRS Documentation Report DR-006. HRS Documentation Report DR-006. Ann Arbor, MI: Survey Research Center, University of Michigan; 2005. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-006.pdf. Accessed April 27, 2016.

[34] Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: The Health and Retirement Study and the Aging, Demographics, and Memory Study. J Gerontol B Psychol Sci Soc Sci. 2011;66 Suppl 1:i162-i71.

[35] Cox DR. Regression models and life-tables. J R Stat Soc B. 1972;34:187-220.

[36] Goodman LA, Kruskal WH. Measures of association for cross classifications. J Am Stat Assoc. 1954;49:732-64.

[37] Herzog AR, Wallace RB. Measures of cognitive functioning in the AHEAD Study. J Gerontol B Psychol Sci Soc Sci. 1997;52 Spec No:37-48.

[38] Freedman VA, Aykan H, Martin LG. Aggregate changes in severe cognitive impairment among older Americans: 1993 and 1998. J Gerontol B Psychol Sci Soc Sci. 2001;56:S100-11.

[39] Suthers K, Kim JK, Crimmins E. Life expectancy with cognitive impairment in the older population of the United States. J Gerontol B Psychol Sci Soc Sci. 2003;58:S179-86.

[40] Lievre A, Alley D, Crimmins EM. Educational differentials in life expectancy with cognitive impairment among the elderly in the United States. J Aging Health. 2008;20:456-77.

[41] Davydow DS, Levine DA, Zivin K, Katon WJ, Langa KM. The association of depression, cognitive impairment without dementia, and dementia with risk of ischemic stroke: a cohort study. Psychosom Med. 2015;77:200-8.

[42] Lin PJ, Kaufer DI, Maciejewski ML, Ganguly R, Paul JE, Biddle AK. An examination of Alzheimer's disease case definitions using Medicare claims and survey data. Alzheimers Dement. 2010;6:334-41.

[43] Yang Z, Zhang K, Lin PJ, Clevenger C, Atherly A. A longitudinal analysis of the lifetime cost of dementia. Health Serv Res. 2012;47:1660-78.

[44] Vu TT, Zhao L, Liu L, Schiman C, Lloyd-Jones DM, Daviglus ML, et al. Favorable cardiovascular health at young and middle ages and dementia in older age-The CHA Study. J Am Heart Assoc. 2019;8:e009730.

[45] Yashkin AP, Akushevich I, Ukraintseva S, Yashin A. The effect of adherence to screening guidelines on the risk of Alzheimer's Disease in elderly individuals newly diagnosed with type 2 diabetes mellitus. Gerontol Geriatr Med. 2018;4:1-9.

[46] Pressley JC, Trott C, Tang M, Durkin M, Stern Y. Dementia in community-dwelling elderly patients: A comparison of survey data, medicare claims, cognitive screening, reported symptoms, and activity limitations. J Clin Epidemiol. 2003;56:896-905.

[47] Amjad H, Roth DL, Sheehan OC, Lyketsos CG, Wolff JL, Samus QM. Underdiagnosis of Dementia: an Observational Study of Patterns in Diagnosis and Awareness in US Older Adults. J Gen Intern Med. 2018;33:1131-38.

FIGURE LEGEND

Figure 1. Schematic summary: Study design and Coding Scheme: The figure shows the study design and the definition of ADRD based on the claims-based algorithm. In an example, it shows a three year look-back period from age 85 for six patients. The appearance of an ADRD code in claims for each patient is marked '|' The shaded horizontal bar shows stay in a nursing home. The scores based on the 4 indicators are shown in the right column. Patients with a score of ≥2 are labeled as ADRD.

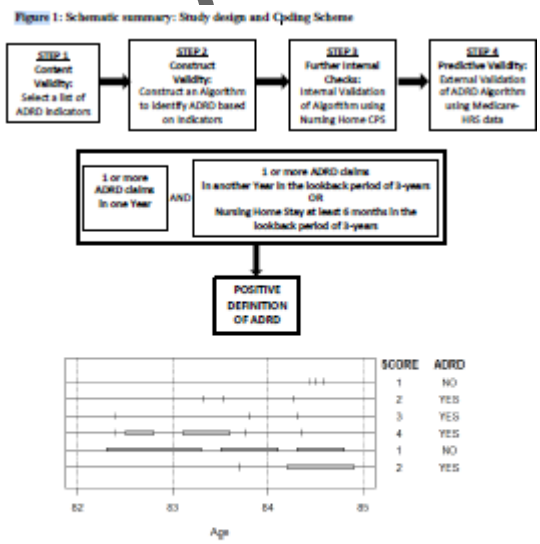

Figure legend: The figure shows the study design and the definition of ADRD based on the claims-based algorithm. In an example, it shows a three-year look-back period from age 85 for six patients. The appearance of an ADRD code in claims for each patient is marked '|' The shaded horizontal bar shows stay in a nursing home. The scores based on the 4 indicators are shown in the right column. Patients with a score of ≥2 are labeled as ADRD.

Table 1. Alzheimer's Disease and Related Disorders	
ICD-9 code	
3310	Alzheimer's disease
2900	Senile dementia, uncomplicated
29010	Presenile dementia, uncomplicated
29011	Presenile dementia with delirium
29012	Presenile dementia with delusional features
29013	Presenile dementia with depressive features
29020	Senile dementia with delusional features
29021	Senile dementia with depressive features
2903	Senile dementia with delirium
29040	Vascular dementia, uncomplicated
29041	Vascular dementia, with delirium
29042	Vascular dementia, with delusions
29043	Vascular dementia, with depressed mood
2940	Amnestic disorders in conditions classified elsewhere
29410	Dementia in conditions classified elsewhere without behavioral disturbance
29411	Dementia in conditions classified elsewhere with behavioral disturbance
29420	Dementia, unspecified, without behavioral disturbance
29421	Dementia, unspecified, with behavioral disturbance
2948	Other persistent mental disorders due to conditions classified elsewhere
797	Senility without mention of psychosis
ICD-10 codes	
G300	Alzheimer's disease with early onset
G301	Alzheimer's disease with late onset
G308	Other Alzheimer's disease
G309	Alzheimer's disease, unspecified
F0150	Vascular dementia without behavioral disturbance
F0151	Vascular dementia with behavioral disturbance
F0280	Dementia in other diseases classified elsewhere without behavioral disturbance
F0281	Dementia in other diseases classified elsewhere with behavioral disturbance
F0390	Unspecified dementia without behavioral disturbance
F0391	Unspecified dementia with behavioral disturbance
F04	Amnestic disorder due to known physiological condition
R4181	Age-related cognitive decline

Table 2. Estimates of getting an ADRD code after 75, 80, and 85 year of age based on codes in the 3-year lookback unadjusted and adjusted for total number of codes

		Unadjusted		Adjusted	
	ADRD Codes	**Hazard ratio**	**p-value**	**Hazard ratio**	**p-value**
Age 75	Single year	-ref-	-ref-	-ref-	-ref-
	Multiple years	4.386	<.0001	2.297	<.0001
	None	0.122	<.0001	0.456	<.0001
	Total codes	--	--	1.035	<.0001
Age 80	Single year	-ref-	-ref-	-ref-	-ref-
	Multiple years	3.806	<.0001	2.482	<.0001
	None	0.154	<.0001	0.432	<.0001
	Total codes	--	--	1.018	<.0001
Age 85	Single year	-ref-	-ref-	-ref-	-ref-
	Multiple years	3.027	<.0001	2.024	<.0001
	None	0.183	<.0001	0.439	<.0001
	Total codes	--	--	1.026	<.0001

Table 2 shows the result from cox proportional hazard model showing ratio of getting an ADRD code after age 75, 80, and 85 who have no codes and codes in multiple years vs codes in single year with and without adjusting for number of codes

Table 3. Odds ratio showing unidimensionality between three separate indicators for ADRD for each year, and an indicator for 6-months stay in a nursing home in a 3-years look back

	ADRD 2-years ago	ADRD 3-years ago	6-months nursing home stay
Age 75			
ADRD 1-year ago	24.549 *	8.165*	3.436*
ADRD 2-years ago	.	14.741*	2.754*
ADRD 3-years ago	.	.	2.764*
Age 80			
ADRD 1-year ago	21.356*	7.854*	3.224*
ADRD 2-years ago	.	12.328*	2.607*
ADRD 3-years ago	.	.	2.369*
Age 85			
ADRD 1-year ago	16.526*	6.409*	2.643*
ADRD 2-years ago	.	11.861*	2.532*
ADRD 3-years ago	.	.	2.287*

* p-value <0.0001

Table 3 shows results from the Mantel-Hanszel test displaying odds of three indicators of ADRD and other conditions for each year, and a 6-months stay in a nursing home in the 3 years lookback period for age 85, 80 and 75 years.

Table 4. Distribution of HRS survey scores* from Self and Proxy respondents stratified by scores from the claims-based coding algorithm

	Self-respondents						Proxy respondents					
Claims-based score	**N**	**(%)**	**HRS Mean score**	**Std. Dev**	**Gamma (95% CI)**	**Probability of Concordance**	**N**	**(%)**	**HRS Mean score**	**Std. Dev**	**Gamma (95% CI)**	**Probability of Concordance**
Age 75												
0-1	3,983	(97.94%	21.81	4.96	-0.68		1776	(70.08%	3.44	2.52	0.78	
≥2	83	(2.06%)	13.93	6.66	(-0.77, -0.58)	83.80%		(29.92%)	7.32	2.07	(0.70, 0.87)	89.24 %
Total	4,037	(100.00%)	21.72	5.10			254	(100.00%)	4.60	2.99		
Age 80												
0-1	3,007	(96.07%	20.71	4.86	-0.62		1613	(56.15%	4.18	2.82	0.67	
≥2	123	(3.93%)	14.10	6.16	(-0.70, -0.54)	81.06%	2	(43.85%)	7.53	2.20	(0.57, 0.77)	83.43%
Total	3,130	(100.00%)	20.53	5.12			301	(100.00%)	5.65	3.06		
Age 85												
0-1	2,009	(89.35%	19.51	5.05	-0.62		1416	(44.23%	4.36	2.52	0.65	
≥2	168	(10.65%)	13.10	5.83	(-0.69, -0.55)	81.04%	6	(55.77%)	7.34	2.44	(0.55, 0.75)	82.63%
Total	2,177	(100.00%)	19.02	5.41			312	(100.00%)	5.95	2.83		

*Self-respondents were scored on a scale of 0-35 where lower scores indicate poor cognition, and proxy respondents were scored on a scale of 0-11 where higher scores indicate poor cognition; Probability shown is the Probability of Concordance

Table 4 shows the mean HRS survey scores from self and proxy respondents for patients who scored ≤1 and ≥2 on the claims-based algorithm. The gammas (γ) are displayed to show the degree of agreement between both scores. The last column shows the probability of concordance estimated by (1+| γ|)/2.

Table 5. Odds ratio between definition of dementia by HRS survey and claims based coding algorithm

Cognitive impairment defined by HRS survey	ADRD defined by claims based coding algorithm: No	Yes	Total (%)	Relative Risk (95 % CI)	Odds Ratio (95 % CI)	p-value
Age 75						
No	4,035 (98.22%)	73 (1.78%)	4,108 (95.74 %)	23.04 (18.07, 29.39)	49.01 (33.81, 71.04)	<0.0001
Yes	97 (53.00%)	86 (47.00%)	183 (4.26 %)			
Total (%)	4,132 (96.29%)	159 (3.71%)	4,291 (100.0			
Age 80						
No	3,064 (96.50%)	111 (3.50%)	3,175 (92.54 %)	16.01 (12.96, 19.78)	35.49 (26.00, 48.45)	<0.0001
Yes	112 (43.75%)	144 (56.25%)	256 (7.46 %)			
Total (%)	3,176 (92.57%)	255 (7.43%)	3,431 (100.0			
Age 85						
No	2,048 (92.75%)	160 (7.25%)	2,208 (88.71 %)	10.49 (8.49, 12.96)	20.81 (15.58, 27.81)	<0.0001
Yes	107 (38.08%)	174 (61.92%)	281 (11.29 %)			
Total	2,155 (86.58%)	334 (13.42%)	2,489 (100.0			

Table 5 compares our ADRD definition based on the claims-based algorithm with the HRS definition of cognitive impairment. The relative risk is the multiplicative increase in the probability of cognitive impairment by the HRS survey predicted by our measure of ADRD. The odds ratio is the multiplicative increase in odds of cognitive impairment by the HRS survey predicted by our measure of ADRD.

RESEARCH IN CONTEXT (149/150)

1. Systematic review: Medicare claims have been widely used to identify Alzheimer's disease and related dementias (ADRD). However, little work has been done to validate claims-based measures in terms of the standard concepts of content, construct, and predictive validity.
2. Interpretation: We propose a new claims-based measure of ADRD that: (i) focuses on ADRD, (ii) requires at least two indicators separated by a significant period of time, (iii) exhibits; (a) content validity in terms of its definition, (b) construct validity in terms of interrelationships of its components, and ability to predict future claims, and (c) predictive validity in its association with two direct measures of cognitive performance from (1) the Health and Retirement Study, and (2) nursing home assessments.
3. Future directions: A valid claims-based measure of ADRD offers the potential to study ADRD in large populations at a limited cost. The limitations of claims-based studies are also discussed.

HIGHLIGHTS (85/85 words)

- Our claims-based definition of ADRD predicts both, claims and non-claims, measures of ADRD and/or cognitive function.
- The four indicators in our claims-based definition of ADRD exhibits a pattern of internal consistency expected from indicators of a single or unidimensional quantity.
- The number of claims for ADRD is less important than claims separated by a significant period of time.
- Individuals with claims over a significant period of time are more likely to have subsequent codes for ADRD than those who have codes close together in time.