Title

Development and validation of machine learning models to predict risk of undiagnosed dementia using healthcare claims and electronic health record data.

Journal

Journal of Alzheimer's Disease

ISSN

1387-2877

Authors

Barnes, Deborah E
Benjamin, Cynthia
Boscardin, W John

Publication Date

2026-08-12

DOI

10.1177/13872877261474534

Development and Validation of Machine Learning Models to Predict Risk of Undiagnosed Dementia Using Healthcare Claims and Electronic Health Record Data

Running title: Machine Learning to Detect Undiagnosed Dementia

Deborah E. **Barnes**, PhD, MPH;[1] Cynthia **Benjamin**;[2] W. John **Boscardin**, PhD[1]

[1]University of California, San Francisco

[2]Together Senior Health, Inc.

Word count: 3,980

Corresponding author:

Deborah Barnes

675 18th Street, Box 0984

San Francisco, CA 94143-0984

Phone: 415-298-5498

Email: Deborah.barnes@ucsf.edu

Abstract

Background: Approximately half of people living with Alzheimer's disease and related dementias are undiagnosed.

Objective: To develop and validate algorithms that predict risk of undiagnosed dementia using electronic health record (EHR) and/or healthcare claims data.

Methods: Study participants were adult patients aged 65 years or older without evidence of dementia (diagnosis/medication) at baseline in two U.S. data sources: 1) Medicare claims (2010 to 2021); 2) EHR and claims from a primary care network (2016 to 2023). We applied coefficients from an existing, validated EHR-based algorithm to predictors defined using Medicare claims and used machine learning to develop new EHR- and claims-based predictive models. We assessed model discrimination using c-statistics.

Results: Study participants included 8,374,400 Medicare beneficiaries (mean [SD] age, 76 [7] years; 57% female) and 29,983 primary care patients (age: 75 [6] years; 56% female). Model discrimination was good when applying EHR-based coefficients to Medicare claims-based predictors (c-statistic [95% confidence interval]: 0.770 [0.767, 0.773]) and was improved by refitting the model (0.795 [0.792, 0.798]) with a small added benefit from incorporating new claims-based predictors (0.801 [0.799; 0.804]). Similarly, when both EHR and claims data were available, discrimination was improved by refitting the model with a small additional benefit from including new predictors, regardless of the data source (EHR, claims, or either).

Conclusions: A validated EHR-based algorithm predicted risk of undiagnosed dementia in Medicare claims with good discrimination. Model accuracy was improved by refitting and, to a lesser extent, by including novel predictors.

Keywords: Alzheimer's disease; dementia; delayed diagnosis; undiagnosed diseases; electronic health record data; administrative claims, healthcare; predictive learning models.

Introduction

Approximately 7 million adults in the U.S. are currently living with Alzheimer's disease and related dementias (ADRD),[1] but only half are diagnosed.[2–6] Most people are diagnosed relatively late in the disease process, often in the context of a healthcare crisis that brings unmanaged symptoms to light. Earlier dementia diagnosis has many potential benefits for patients, families, clinicians, and healthcare systems.[7–10] Patients can be assessed for reversible causes[11] and eligibility for newer treatments, which are only approved for earlier-stage disease.[12,13] Patients also can be involved in medical decision-making and planning for the future while they retain capacity.[10] Caregivers can be provided with education and support, which can improve caregiver well-being.[14,15] Clinicians can implement care plans to help patients maximize their quality of life and health,[16–19] potentially reducing the risk of avoidable hospitalizations[20] and risky behaviors such as driving.[21,22] Healthcare systems may experience lower costs and higher patient satisfaction due to better care management.[23,24] Lack of time and competing priorities during primary care visits are often cited as key barriers to recognizing and diagnosing dementia in the healthcare setting.[3,8,25–27]

To support earlier dementia detection and diagnosis and help address time constraints, Barnes et al. developed an algorithm that uses structured electronic health record (EHR) data to identify individuals at greatest risk of having undiagnosed dementia.[28] This algorithm—called the electronic health record Risk of Alzheimer's and Dementia Assessment Rule (eRADAR)—has been validated in two healthcare systems and found to have high prognostic accuracy, with c-statistics ranging from 0.79 to 0.84.[29] In addition, the algorithm demonstrated consistent accuracy over time and in different races, ethnicities, and genders.

A limitation of eRADAR and other EHR-based models is that they typically are not evaluated using other data sources. For example, many insurers and researchers have access to healthcare claims data, rather than EHR data. Some settings have access to both EHR and claims data, but it is not known whether these data sources are interchangeable. It also is not known whether updating coefficients or incorporating new predictors results in meaningful prognostic improvement. Therefore, the primary goals of this study were to 1) determine the accuracy of the eRADAR model when applied in new data sources and 2) assess the impact of new predictors and coefficient calculations on model accuracy.

Methods

Claims-Based Data Source

We used healthcare claims data from the Center for Medicare and Medicaid Services (CMS) 5% samples from 2010 to 2021. We created 10 unique cohorts that each included beneficiaries with three consecutive years of data (i.e., 2010 to 2012, 2011 to 2013, … 2019 to 2021). Within each cohort, data from the first two years were used to identify predictors of undiagnosed dementia, and the third year was used to determine the outcome of incident dementia diagnosis, with the assumption that those who received an incident dementia diagnosis during the calendar year were undiagnosed on January 1 of that year.[29] Beneficiaries were included in a given cohort if they were: age 65 years or older for the entire three year period; enrolled for the entire three year period; had at least two claims from a primary care physician (PCP), specialist, urgent care, or emergency department visit during the two-year prediction period; and did not have hospice care

or a dementia diagnosis during the prediction period. Beneficiaries could be included in multiple cohorts if they met all eligibility criteria.

Combined Data Source

A national value-based healthcare partner who cares for over 7 million patients and supports 2800 providers in 26 states provided a dataset that included patients with three years of continuous data for both EHR and administrative claims from 2016 through 2023. For patients with more than three years of data, the most recent three years were provided. Similar to the claims-only analysis, measures from the first 2 years were used to identify predictors of undiagnosed dementia and the third year was used to determine the outcome of incident dementia diagnosis. Patients were excluded if they were not continuously enrolled with our partner for at least three consecutive years, were not 65 or older for the entire study period, had prevalent dementia or received hospice care during the two-year prediction period, or died during the study period. We created a single dataset of 29,983 eligible members with data on demographics (age and gender), vital signs (height, weight, and blood pressure), utilization (service date, current procedural terminology [CPT]/healthcare common procedure coding system [HCPCS] code), prescription drugs, and diagnoses from both EHR and claims sources for analyses.

Variable Harmonization

For each variable in the original EHR-based eRADAR model, we created a harmonized variable using claims data. In many cases, we used the same International Classification of Diseases-9th or 10th edition (ICD-9/10) codes that were used in eRADAR. For variables that were not originally

114 defined based on ICD codes, we developed claims-based proxies using existing literature and
115 expert consultation. For example, we used ICD codes rather than vital signs to define obesity,
116 underweight, and high blood pressure. Similarly, we used ICD codes rather than medication fills
117 or prescriptions as proxies for use of anti-depressants or sleep medications. For healthcare
118 utilization variables, we used HCPCS codes or Milliman's Health Cost Guidelines™, a
119 proprietary software model that sorts medical and pharmacy claims data into hospital, surgical,
120 medical, and other benefit service categories using Health Cost Guidelines (HCG) definitions. In
121 addition, we created new claims-based predictor variables based on existing literature (e.g.,
122 Lancet Commission Report[30]) and internal team discussions. The complete list of predictor
123 variables and their definitions in EHR and claims data is included in Supplemental Table 1.
124
125 *Analytic Approach*
126 *Claims-Based Model Development and Validation.* For each of the 10 CMS cohorts, we fit six
127 different models and compared how well they predicted the likelihood that a beneficiary was
128 diagnosed with incident dementia during the year (i.e., had undiagnosed dementia at the
129 beginning of the follow-up year). Model 1: Coefficients from the eRADAR model were applied
130 to harmonized variables from claims data. Model 2: Logistic regression was performed using the
131 Model 1 variables to calculate new coefficient values. Model 3: New claims-based predictors
132 were combined with Model 1 predictors, and Least Absolute Shrinkage and Selection Operator
133 (LASSO) regression was used for variable selection with the lambda value that provided the
134 smallest cross-validated prediction error (LASSO min). Model 4: Same as Model 3, except
135 LASSO used the largest lambda value whose error was within one standard deviation (SD) of the
136 minimum error for variable selection (LASSO 1se). Model 5: These models used LASSO min

but required all the predictors from Model 1 to be included in the final model (i.e., the variable selection was only performed on the new predictors) (Restricted LASSO min). Model 6: These models used LASSO 1se but required all the predictors from Model 1 to be included in the final model (Restricted LASSO 1se).

To minimize the potential for over-fitting, we developed each model using data from one cohort and then applied the results to data from the subsequent cohort to validate model accuracy. For example, to predict the probability of a member having undiagnosed dementia on January 1, 2015, we fit each model using data from the 2014 cohort (i.e., predictors from 2012 and 2013 and outcomes from 2014) and applied the results to the 2015 cohort (i.e., predictors from 2013 and 2014 and outcomes from 2015).

We compared discrimination of the models using c-statistics, also known as area under the receiver operating characteristic curve (AUC or ROC). We also calculated the c-statistics separately by race/ethnicity and sex/gender to determine whether model performance differed based on demographic characteristics. We calculated sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) as well as the actual and predicted probabilities of undiagnosed dementia at various thresholds and assessed calibration by plotting the proportion of incident dementia diagnoses by decile of predicted risk.

Comparison of EHR- and Claims-Based Models. To assess the impact of data source (EHR and/or claims) on model accuracy, we selected the best-fitting model from the most recent Medicare cohort ("optimal Medicare model") and tested it using the combined data source under

different scenarios. For each predictor in the selected model, we determined if it was present in EHR, claims, or either data source during the two-year prediction period. We summed the coefficients for predictors identified as present under each scenario to calculate risk scores. We then determined which patients were newly diagnosed with dementia during the outcome year. New dementia diagnosis could be in either EHR or claims so that each data source was predicting the same outcome. We also refit logistic regression models to calculate new coefficients for predictors from the optimal Medicare model separately for each data source. Finally, we refit logistic regression models considering the full list of potential predictors separately for EHR, claims, or either data source. We calculated c-statistics to assess overall model accuracy under each scenario.

Relative Importance of Variables Included. To help clarify which factors were driving predictions and the relative importance of specific predictors, we calculated population attributable fractions (PAFs)[31,32] using the combined data source with predictors from the optimal Medicare model and refit coefficients. PAFs combine information on the prevalence of a given risk factor with the strength of its association with the outcome to estimate its relative importance. PAFs can be interpreted as estimating the hypothetical relative change in outcome prevalence (i.e., undiagnosed dementia) at the population level if the risk factor were to be eliminated (i.e., totally absent vs. present at the observed rate), with all other factors as observed. For example, elimination of a risk factor with a PAF of 20% would be hypothesized to result in a 20% relative reduction in undiagnosed dementia in the population (e.g., from 10% to 8%).

IRB Approval

All study procedures were approved by the Human Research Protection Program at the University of California, San Francisco (UCSF) (23-38575).

Role of the Funding Source

This study was funded by a Small Business Innovation Research grant from the National Institutes of Aging at the National Institutes of Health (R44AG082605). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funder was not involved with the study design; collection, analysis, or interpretation of data; or the decision to approve publication of the finished manuscript.

Results

Claims-Based Participants

The ten Medicare 5% sample cohorts had a mean (SD) number of 1,511,989 (40,456) unique patients, and the total across the cohorts was 15,121,141 patients (Supplemental Table 2). Of these, 6,746,725 (45%) were excluded because they did not have 3 consecutive years of data (18%), were less than 65 years old (16%), did not have adequate claims data during the 2-year predictor period (6%) or had claims for dementia or hospice during the predictor period (6%) leaving a final sample size of 8,374,400 eligible beneficiaries. Cohort members had a mean (SD) age of 76 (7) years; 57% were female, and the race/ethnicity distribution was 79% non-Hispanic white, 6% African American/black; 2% Asian; 1% Hispanic/Latino; 0.4% North American Native; and 13% another or unknown race/ethnicity (Table 1). The five most common comorbid conditions were hypertension (80%), hypercholesterolemia (79%), diabetes (32%), chronic pulmonary disease (27%), and hypothyroidism (26%). Approximately one-third (32%) had an

emergency department visit during the two-year predictor period. The prevalence of all predictors is provided in Supplemental Table 3.

Combined Data Source Participants

A total of 29,983 unique patients in the combined data source cohort had three consecutive years of both EHR and claims data (Table 1). They had a mean (SD) age of 75 (6) years, and 56% were female. Race/ethnicity data were not provided. We observed several differences in the prevalence of some predictor variables in the combined cohort compared to the Medicare cohort (Table 1, Supplemental Table 3). These included anti-depressants (54% vs. 15%), sleep aids (53% vs. 14%), tobacco use (38% vs. 23%), renal disease (27% vs. 16%), peripheral vascular disease (23% vs. 11%), and obesity (16% vs. 5%). Prevalence of most other predictor variables was comparable in the two data sources.

Accuracy of Claims-Based Models

Model 1, in which the original eRADAR coefficients were applied to predictor variables redefined using claims, yielded an average c-statistic of 0.770 (95% confidence interval [CI]: 0.767, 0.773) over the cohorts (outcomes from 2013 to 2021). Model 2, in which new coefficients were calculated for the Model 1 predictors, resulted in improvement in accuracy (average c-statistic: 0.795; 95% CI: 0.792, 0.798). Small additional improvements in c-statistics were observed in Models 3 to 6, which explored the impact of including new predictor variables under various scenarios (Table 2). The highest c-statistics were obtained using LASSO min, which included new predictors and updated coefficients without restrictions (average: 0.801 [0.799, 0.804]). Full model coefficients for the original eRADAR model and LASSO Min 2021

(optimal Medicare model) are provided in Supplemental Table 4. Average c-statistics were similar for female and male beneficiaries but were slightly lower for black and Hispanic than white beneficiaries (Figure 1).

In the LASSO 1se models, in which parsimony was prioritized over discrimination, many of the original eRADAR predictors were dropped from the model (Supplemental Table 4). These included gender, chronic pulmonary disease, hypothyroidism, renal failure, lymphoma, rheumatoid arthritis, blood loss anemia, tobacco use, atrial fibrillation, high BMI, high blood pressure, and sleep medication. When the original predictors were forced to remain in the model (Models 5 and 6), the only new variables that were always added were mild cognitive impairment (MCI) and Parkinson's disease. When all potential predictors were allowed to compete with each other for inclusion (Models 3 and 4), additional variables that were always selected included delirium, peripheral vascular disease, seizure, and mental health visit. Sensitivity and PPV were slightly better in refit models, while specificity and NPV were similar across all models (Table 2).

The calibration of the LASSO min model over decile of predicted risk averaged over the 10 Medicare cohorts is shown in Supplemental Figure 1. Predicted risk of undiagnosed dementia using data from the two-year prediction period increased in a linear manner from <1% in the lowest decile to approximately 13% in the highest decile, and these values were comparable to actual rates of undiagnosed dementia observed during the outcome period.

Sensitivity for accurately identifying undiagnosed dementia at different risk score thresholds using data from the optimal Medicare model is shown in Figure 2. We estimated that targeted screening of patients with risk scores in the top 15% could potentially identify 54% of patients with undiagnosed dementia. Targeted screening of patients with risk scores in the top 35% could potentially identify nearly 80% of those with undiagnosed dementia.

Comparison of Claims and EHR Data Sources

When we tested the optimal Medicare model in the combined data source with both EHR and claims data, we found that discrimination was slightly lower using EHR data (0.746 [0.731, 0.762]) compared to claims data (0.763 [0.748, 0.777]) or both data sources (0.765 [0.750, 0.779]). Discrimination was improved by refitting the model using the same predictors, with little difference based on data source (EHR: 0.766 [0.751, 0.780]); claims: 0.773 [0.759, 0.787]; either: 0.777 [0.762, 0.791]). Refitting the model using the full list of available predictors resulted in negligible additional improvements, regardless of the data source (EHR: 0.767 [0.752, 0.781]; claims: 0.776 [0.762, 0.790]; either: 0.778 [0.764,0.792]) (Figure 3).

Relative Importance of Variables Included

Our analysis of PAFs revealed that the variables associated with the largest proportion of undiagnosed dementia included hypertension (17%), depression (11%)/antidepressants (6%), cerebrovascular disease (10%), peripheral vascular disease (8%), MCI (7%), and diabetes (5%). For comparison, the PAF for age was 8% per year. All other predictors had PAFs of <5% (Supplement Table 6).

Discussion

This study of more than 8 million Medicare beneficiaries and nearly 30,000 patients from a national primary care network found that both EHR and claims data can be used to predict a patient's risk of undiagnosed dementia with accuracy comparable to many other prognostic algorithms such as those used to predict risk of cardiovascular disease,[33,34] breast cancer,[35,36] mortality,[37,38] and other health conditions.[39] Application of coefficients from the publicly available, validated, EHR-based eRADAR algorithm[28,29] to predictors defined using claims data yielded good prognostic accuracy, suggesting that the original algorithm is valid using either EHR or claims data.

We also found that model accuracy was improved by refitting the data for a given population. This was true for both the Medicare cohorts and the primary care cohort and likely reflects differences in patient characteristics between populations that result in model improvements with slight reweighting of core predictors. We do not believe this improvement is due to over-fitting because, for each of the Medicare samples, we developed models in one cohort and then tested their accuracy (i.e., validated) in the subsequent cohort. Users seeking maximum accuracy for predicting undiagnosed dementia may choose to recalculate model coefficients using their own data. However, we note that this requires three years of consecutive data on a sample of at least 10,000 older patients without prevalent dementia diagnoses as well as the ability to define the key predictors and run LASSO regression analyses.

We were surprised that incorporation of new claims-based predictors had relatively little impact on model accuracy. This may reflect the comprehensiveness of the original list of predictors

297 considered for the eRADAR algorithm. The predictors that were most consistently included in
298 new models were diagnoses of MCI and Parkinson's disease, both of which are consistent with
299 research identifying them as risk factors for dementia.[40–43] People with MCI were excluded from
300 the original eRADAR development and validation studies, but we decided to include these
301 individuals in the present study since MCI is often a precursor to developing dementia.[40]
302 Parkinson's has been previously identified as a variable that should be considered for inclusion in
303 future models.[28] Most new predictors identified for inclusion were relatively rare in these cohorts,
304 so their inclusion did not have a large impact on overall model accuracy.
305
306 Although our modeling approach was predictive rather than etiologic, our PAF analysis provides
307 some indication of the relative importance of various predictors. Factors with higher PAFs
308 included hypertension, depression, cerebrovascular disease, peripheral vascular disease, MCI,
309 and diabetes diagnoses. All these factors have been associated with increased dementia risk in
310 prior studies, although MCI diagnoses may also reflect clinical awareness of existing cognitive
311 decline.[44–48] Future studies should assess the impact of interventions to reduce or address risk
312 factors, particularly those that are modifiable.
313
314 Numerous algorithms have been developed to predict risk of future dementia. A recently updated
315 systematic review identified 74 dementia risk predictions models,[49] of which more than 50 had
316 been published since 2014. However, most include predictors that are not always readily
317 available (e.g., education, apolipoprotein-E genotype), few were developed in population-
318 representative samples, and few have been externally validated. Twelve used EHR or claims
319 data, but none of these were both representative and externally validated. The authors concluded

that there is an urgent need to develop a robust, validated prediction model in the general population that can be widely implemented in clinical practice to improve dementia detection and prevention. The algorithm described in this manuscript may address this need.

Another recent systematic review identified "passive digital markers" for dementia, which are EHR-based algorithms that utilize existing healthcare data to support earlier detection and diagnosis.[50] A total of 19 passive digital markers for dementia were identified, most of which were rated as having poor quality based on standard criteria. eRADAR was one of the highest-rated algorithms in this review. Our current study validates eRADAR for use in claims data and also identifies strategies for improving its accuracy.

This manuscript has several key strengths. First, we are unaware of any other validated predictive models that use claims data to identify patients at risk of having undiagnosed dementia. Second, our models were developed and validated using two large, national datasets. The Medicare 5% samples are nationally representative of older adults who receive healthcare in the U.S., and our analyses utilized 10 unique cohorts comprising more than 8 million individuals. Our primary care partner currently serves over 7 million patients in 26 states, and our analytic dataset included nearly 30,000 patients. Finally, we compared models within each cohort to assess the added value of recalculating model coefficients, including new predictors, and utilizing different data sources. Our finding that the greatest improvements in model accuracy are achieved by refitting the model in a given data source, with little additional value of adding new predictors or data sources, may help others who are interested in implementing a tool to support detection of undiagnosed dementia in their healthcare systems or networks.

Our study also has limitations. Perhaps most importantly, this study illustrates clearly that there is unlikely to be a single model that fits all needs. In fact, we fit dozens of models that had similar discriminative ability over the course of performing this study. We also did not explore the impact of using other machine learning or artificial intelligence techniques, although our prior studies have found that there is typically relatively little gain in accuracy with these techniques compared to LASSO.[44,45] Although our sample sizes were large, our Medicare sample had smaller numbers of Asian, Black, Hispanic, and North American Native individuals resulting in less precise estimates of model accuracy, and we did not have data on race or ethnicity in our combined data source sample. Future studies should consider reweighting or subgroup-specific calibration to maximize model accuracy in different populations.

We chose to only consider new predictors with biological plausibility and to focus on structured rather than unstructured data elements. It is possible that inclusion of other predictors including free text could have improved prognostic ability; however, this would come at a cost of interpretability and transportability. A key advantage of our approach is the ability to apply the results in other settings with relative ease. We also recognize that some of the predictors (e.g., MCI, Parkinson's disease) may identify individuals who are already known by clinicians to be at risk and not truly "unrecognized." None-the-less, we believe that identification of these individuals could serve as a reminder to clinicians to follow up if needed, particularly given that treatments are more likely to be effective if started earlier in the disease process. Finally, a risk score alone is unlikely to change clinical practice. Additional studies are needed to determine the optimal strategies for implementing dementia risk algorithms in healthcare settings and sharing

risk prediction results with healthcare providers and patients to support earlier detection and diagnosis.

Payers – both Medicare and private insurers – are increasingly interested in early detection of cognitive decline, but the time and resource challenges of broad screening remain. With growing accuracy of digital diagnostics and blood-based biomarkers,[51] drugs such as lecanemab and donanemab that are showing efficacy for some early-stage patients,[52,53] and lifestyle-based interventions that are proving to be effective,[54,55] the financial value of early intervention is becoming more apparent.[56] A claims-based tool that accurately identifies patients at increased risk of undiagnosed dementia could enable new implementation models designed to identify high-risk individuals in the background without increasing the immediate burden on primary care physicians. For example, payers could initiate direct outreach to patients to provide education, address potential fears and concerns about stigma, and sensitively encourage those at highest risk to seek cognitive screening and treatment. High-risk patients also could also be provided with access to targeted care management to address risk factors, provide diagnoses, and manage symptoms.

In conclusion, this study demonstrated that both EHR and healthcare claims data can be utilized to create risk scores for undiagnosed dementia that have good to excellent prognostic accuracy. These models can be optimized by recalculating coefficients using local data; however, accuracy without refitting is still comparable to other commonly used prognostic models. Healthcare plans and provider networks could potentially utilize this approach to support earlier detection and diagnosis of dementia in their patients.

Acknowledgements

We would like to thank Delia Mannen for her work preparing data for these analyses. We also thank Spencer Marshall, Christopher Kunkel, and Brian Hartman at Milliman for preparing data, performing analyses, and publishing a white paper available at: https://www.milliman.com/en/insight/analysis-of-together-senior-healths-radar-algorithm.

Author contributions

Dr. Barnes contributed to conception and design of the work, interpretation of data for the work, obtaining funding, and drafting the manuscript.

Ms. Benjamin contributed to conception and design of the work, acquisition and interpretation of data for the work, obtaining funding, and reviewing the manuscript critically for important intellectual content.

Dr. Boscardin contributed to design of the work, analysis and interpretation of data for the work, and reviewing the manuscript critically for important intellectual content.

Statement and declarations

Ethical considerations

All study procedures were approved by the Human Research Protection Program Institutional Review Board (IRB) at the University of California, San Francisco (IRB#: 23-38575).

Consent to participate

A waiver of consent was approved by the IRB because the research posed no more than minimal risks and could not practicably be done without the waiver.

Consent for publication

Not applicable.

Declaration of conflicting interest

Dr. Barnes and Ms. Benjamin are co-founders of Together Senior Health, a wholly owned subsidiary of Linus Health, and have stock options. Linus Health has developed proprietary versions of the algorithms described in this manuscript. Conflicts of interest were reviewed and managed by the Conflict of Interest Advisory Committee at UCSF. Dr. Boscardin has no conflicts of interest and served as the independent biostatistician for this study. He supervised all analyses and independently performed analyses in the combined data source.

Funding statement

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (R44AG082605). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability

Data underlying this work are governed by Data Use Agreements that do not allow for sharing with other investigators.

Supplemental material

Submitted separately.

References

1. 2024 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2024;20(5):3708-3821. doi:10.1002/alz.13809

2. Lang L, Clifford A, Wei L, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open*. 2017;7(2):e011146. doi:10.1136/bmjopen-2016-011146

3. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and Delayed Diagnosis of Dementia in Primary Care: Prevalence and Contributing Factors. *Alzheimer Disease & Associated Disorders*. 2009;23(4):306-314. doi:10.1097/WAD.0b013e3181a6bebc

4. Boise L, Neal MB, Kaye J. Dementia Assessment in Primary Care: Results From a Study in Three Managed Care Systems. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2004;59(6):M621-M626. doi:10.1093/gerona/59.6.M621

5. Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Underdiagnosis of dementia in primary care: Variations in the observed prevalence and comparisons to the expected prevalence. *Aging & Mental Health*. 2011;15(8):978-984. doi:10.1080/13607863.2011.596805

6. Ganguli M, Rodriguez E, Mulsant B, et al. Detection and Management of Cognitive Impairment in Primary Care: The Steel Valley Seniors Survey. *J American Geriatrics Society*. 2004;52(10):1668-1675. doi:10.1111/j.1532-5415.2004.52459.x

7. Borson S, Frank L, Bayley PJ, et al. Improving dementia care: The role of screening and detection of cognitive impairment. *Alzheimer's & Dementia*. 2013;9(2):151-159. doi:10.1016/j.jalz.2012.08.008

8. Connell CM, Boise L, Stuckey JC, Holmes SB, Hudson ML. Attitudes Toward the Diagnosis and Disclosure of Dementia Among Family Caregivers and Primary Care Physicians. *The Gerontologist*. 2004;44(4):500-507. doi:10.1093/geront/44.4.500

9. Dubois B, Padovani A, Scheltens P, Rossi A, Dell'Agnello G. Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. Saykin A, ed. *JAD*. 2015;49(3):617-631. doi:10.3233/JAD-150692

10. Power MC, Willens V, Prather C, et al. Risks and Benefits of Clinical Diagnosis Around the Time of Dementia Onset. *Gerontol Geriatr Med*. 2023;9:23337214231213185. doi:10.1177/23337214231213185

11. Piccini C, Bracco L, Amaducci L. Treatable and reversible dementias: an update. *J Neurol Sci*. 1998;153(2):172-181. doi:10.1016/s0022-510x(97)00289-x

12. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *JAMA*. 2023;330(6):512. doi:10.1001/jama.2023.13239

13. Van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in Early Alzheimer's Disease. *N Engl J Med*. 2023;388(1):9-21. doi:10.1056/NEJMoa2212948

14. Huis In Het Veld JG, Verkaik R, Mistiaen P, van Meijel B, Francke AL. The effectiveness of interventions in supporting self-management of informal caregivers of people with dementia; a systematic meta review. *BMC Geriatr*. 2015;15:147. doi:10.1186/s12877-015-0145-6

15. Parker D, Mills S, Abbey J. Effectiveness of interventions that assist caregivers to support people with dementia living in the community: a systematic review. *Int J Evid Based Healthc*. 2008;6(2):137-172. doi:10.1111/j.1744-1609.2008.00090.x

16. 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*. 2006;295(18):2148-2157. doi:10.1001/jama.295.18.2148

17. Dreier-Wolfgramm A, Michalowsky B, Austrom MG, et al. Dementia care management in primary care : Current collaborative care models and the case for interprofessional education. *Z Gerontol Geriatr*. 2017;50(Suppl 2):68-77. doi:10.1007/s00391-017-1220-8

18. Reuben DB, Tan ZS, Romero T, Wenger NS, Keeler E, Jennings LA. Patient and Caregiver Benefit From a Comprehensive Dementia Care Program: 1-Year Results From the UCLA Alzheimer's and Dementia Care Program. *J Am Geriatr Soc*. 2019;67(11):2267-2273. doi:10.1111/jgs.16085

19. Possin KL, Dulaney S, Sideman AB, et al. Long-term effects of collaborative dementia care on quality of life and caregiver well-being. *Alzheimers Dement*. 2025;21(1):e14370. doi:10.1002/alz.14370

20. Lin PJ, Fillit HM, Cohen JT, Neumann PJ. Potentially avoidable hospitalizations among Medicare beneficiaries with Alzheimer's disease and related disorders. *Alzheimers Dement*. 2013;9(1):30-38. doi:10.1016/j.jalz.2012.11.002

21. Meuleners LB, Ng J, Chow K, Stevenson M. Motor Vehicle Crashes and Dementia: A Population-Based Study. *J Am Geriatr Soc*. 2016;64(5):1039-1045. doi:10.1111/jgs.14109

22. Amjad H, Roth DL, Samus QM, Yasar S, Wolff JL. Potentially Unsafe Activities and Living Conditions of Older Adults with Dementia. *J Am Geriatr Soc*. 2016;64(6):1223-1232. doi:10.1111/jgs.14164

23. Michalowsky B, Blotenberg I, Platen M, et al. Clinical Outcomes and Cost-Effectiveness of Collaborative Dementia Care: A Secondary Analysis of a Cluster Randomized Clinical Trial. *JAMA Netw Open*. 2024;7(7):e2419282. doi:10.1001/jamanetworkopen.2024.19282

24. Guterman EL, Kiekhofer RE, Wood AJ, et al. Care Ecosystem Collaborative Model and Health Care Costs in Medicare Beneficiaries With Dementia: A Secondary Analysis of a Randomized Clinical Trial. *JAMA Intern Med*. 2023;183(11):1222-1228. doi:10.1001/jamainternmed.2023.4764

25. Hinton L, Franz CE, Reddy G, Flores Y, Kravitz RL, Barker JC. Practice Constraints, Behavioral Problems, and Dementia Care: Primary Care Physicians' Perspectives. *J GEN INTERN MED*. 2007;22(11):1487-1492. doi:10.1007/s11606-007-0317-y

26. Hansen EC, Hughes C, Routley G, Robinson AL. General practitioners' experiences and understandings of diagnosing dementia: Factors impacting on early diagnosis. *Social Science & Medicine*. 2008;67(11):1776-1783. doi:10.1016/j.socscimed.2008.09.020

27. Iliffe S, Manthorpe J, Eden A. Sooner or later? Issues in the early diagnosis of dementia in general practice: a qualitative study. *Family Practice*. 2003;20(4):376-381. doi:10.1093/fampra/cmg407

28. Barnes DE, Zhou J, Walker RL, et al. Development and Validation of eRADAR: A Tool Using EHR Data to Detect Unrecognized Dementia. *J American Geriatrics Society*. 2020;68(1):103-111. doi:10.1111/jgs.16182

29. Coley RY, Smith JJ, Karliner L, et al. External Validation of the eRADAR Risk Score for Detecting Undiagnosed Dementia in Two Real-World Healthcare Systems. *J GEN INTERN MED*. 2023;38(2):351-360. doi:10.1007/s11606-022-07736-6

30. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*. 2024;404(10452):572-628. doi:10.1016/S0140-6736(24)01296-0

31. Steenland K, Armstrong B. An Overview of Methods for Calculating the Burden of Disease Due to Specific Risk Factors: *Epidemiology*. 2006;17(5):512-519. doi:10.1097/01.ede.0000229155.05644.43

32. World Health Organization. *Environmental Health: Estimations of Attributable Burden of Disease Due to a Risk Factor*. 2018.

33. D'Agostino RB, Vasan RS, Pencina MJ, et al. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation*. 2008;117(6):743-753. doi:10.1161/CIRCULATIONAHA.107.699579

34. D'Agostino RB, Grundy S, Sullivan LM, Wilson P, for the CHD Risk Prediction Group. Validation of the Framingham Coronary Heart Disease Prediction Scores: Results of a Multiple Ethnic Groups Investigation. *JAMA*. 2001;286(2):180. doi:10.1001/jama.286.2.180

35. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81(24):1879-1886. doi:10.1093/jnci/81.24.1879

36. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst*. 1999;91(18):1541-1548. doi:10.1093/jnci/91.18.1541

37. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *J Gen Intern Med*. 2009;24(10):1115-1122. doi:10.1007/s11606-009-1073-y

38. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA*. 2006;295(7):801-808. doi:10.1001/jama.295.7.801

39. Abdulazeem H, Whitelaw S, Schauberger G, Klug SJ. A systematic review of clinical health conditions predicted by machine learning diagnostic and prognostic models trained or validated using real-world primary health care data. Vathy-Fogarassy Á, ed. *PLoS ONE*. 2023;18(9):e0274276. doi:10.1371/journal.pone.0274276

40. Petersen RC, Roberts RO, Knopman DS, et al. Mild cognitive impairment: ten years later. *Arch Neurol*. 2009;66(12):1447-1455. doi:10.1001/archneurol.2009.266

41. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med*. 2008;148(6):427-434. doi:10.7326/0003-4819-148-6-200803180-00005

42. Gallagher J, Gochanour C, Caspell-Garcia C, et al. Long-Term Dementia Risk in Parkinson Disease. *Neurology*. 2024;103(5):e209699. doi:10.1212/WNL.0000000000209699

43. Åström DO, Simonsen J, Raket LL, et al. High risk of developing dementia in Parkinson's disease: a Swedish registry-based study. *Sci Rep*. 2022;12(1):16759. doi:10.1038/s41598-022-21093-8

44. Parfenov VA, Zakharov VV, Kabaeva AR, Vakhnina NV. Subjective cognitive decline as a predictor of future cognitive decline: a systematic review. *Dement neuropsychol*. 2020;14(3):248-257. doi:10.1590/1980-57642020dn14-030007

45. Boparai JK, Clemens M, Jat K. Late-Onset Depression and Dementia: A Systematic Review of the Temporal Relationships and Predictive Associations. *Medicina*. 2025;61(5):839. doi:10.3390/medicina61050839

46. Cao F, Yang F, Li J, et al. The relationship between diabetes and the dementia risk: a meta-analysis. *Diabetol Metab Syndr*. 2024;16(1):101. doi:10.1186/s13098-024-01346-4

47. Salemme S, Lombardo FL, Lacorte E, et al. The prognosis of mild cognitive impairment: A systematic review and meta-analysis. *Alz & Dem Diag Ass & Dis Mo*. 2025;17(1):e70074. doi:10.1002/dad2.70074

48. Wahl D, Clayton ZS. Peripheral vascular dysfunction and the aging brain. *Aging*. 2024;16(10):9280-9302. doi:10.18632/aging.205877

49. Brain J, Kafadar AH, Errington L, et al. What's New in Dementia Risk Prediction Modelling? An Updated Systematic Review. *Dement Geriatr Cogn Dis Extra*. 2024;14(1):49-74. doi:10.1159/000539744

50. Taylor B, Barboi C, Boustani M. Passive digital markers for Alzheimer's disease and other related dementias: A systematic evidence review. *J Am Geriatr Soc*. 2023;71(9):2966-2974. doi:10.1111/jgs.18426

51. Malek-Ahmadi M, Sharma S, Stipho F, et al. Plasma Phosphorylated Tau 217 and Amyloid Burden in Older Adults Without Cognitive Impairment: A Meta-Analysis. *JAMA Neurol*. 2026;83(1):13. doi:10.1001/jamaneurol.2025.4721

52. Abdelazim K, Allam AA, Afifi B, Abdulazeem H, Elbehiry AI. The efficacy and safety of lecanemab 10 mg/kg biweekly compared to a placebo in patients with Alzheimer's disease: a systematic review and meta-analysis of randomized controlled trials. *Neurol Sci*. 2024;45(8):3583-3597. doi:10.1007/s10072-024-07477-w

53. Lu M, Kim MJ, Collins EC, et al. Posttreatment Amyloid Levels and Clinical Outcomes Following Donanemab for Early Symptomatic Alzheimer Disease: A Secondary Analysis of the TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *JAMA Neurol*. 2025;82(12):1251. doi:10.1001/jamaneurol.2025.3869

54. Baker LD, Espeland MA, Whitmer RA, et al. Structured vs Self-Guided Multidomain Lifestyle Interventions for Global Cognitive Function: The US POINTER Randomized Clinical Trial. *JAMA*. 2025;334(8):681. doi:10.1001/jama.2025.12923

55. Barnes DE, Jiang F, Benjamin C, et al. Livestream, group movement program for people living with cognitive impairment and care partners: A randomized clinical trial. *A&D Transl Res & Clin Interv*. 2024;10(2):e12467. doi:10.1002/trc2.12467

56. 2025 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2025;21(4):e70235. doi:10.1002/alz.70235

44. Austin PC, Harrell FE Jr, Steyerberg EW. Predictive performance of machine and statistical learning methods: Impact of data-generating processes on external validity in the "large N, small p" setting. Stat Methods Med Res. 2021 Jun;30(6):1465-1483. doi: 10.1177/09622802211002867. Epub 2021 Apr 13. PMID: 33848231; PMCID: PMC8188999.

45. Jing B, Boscardin WJ, Deardorff WJ, Jeon SY, Lee AK, Donovan AL, Lee SJ. Comparing Machine Learning to Regression Methods for Mortality Prediction Using Veterans Affairs Electronic Health Record Clinical Data. Med Care. 2022 Jun 1;60(6):470-479. doi: 10.1097/MLR.0000000000001720. Epub 2022 Mar 30. PMID: 35352701; PMCID: PMC9106858.

615 **Table 1. Characteristics of Claims-Based and Combined Data Source Cohorts***

Characteristic	**Claims-Based Cohorts (N=8,374,400)†**	**Combined Data Source Cohort (N=29,983)‡**
Demographics		
Age, years	75.9 (6.9)	75.0 (6.2)
Sex/gender, female	4,637,916 (57.1)	16,786 (56.0)
Race/ethnicity		Not provided
Asian	126,252 (1.5)	
Black	485,978 (5.8)	
Hispanic/Latino	84,011 (1.0)	
North American Native	33,535 (0.4)	
White	6,591,751 (78.7)	
Unknown or another race/ethnicity	1,052,874 (12.6)	
Most Prevalent Predictors§		
Hypercholesterolemia	6,595,026 (78.8)	26,261 (87.6)
Hypertension	6,706,035 (80.1)	24,357 (81.2)
Anti-depressants	1,289,328 (15.4)	16,103 (53.7)
Sleep medications	1,204,285 (14.4)	15,840 (52.8)
Tobacco use	1,893,052 (22.6)	11,506 (38.4)
Emergency department visit	2,723,885 (32.5)	9,036 (30.1)
Diabetes	2,666,579 (31.8)	8,264 (27.6)
Renal disease	1,347,472 (16.1)	8,062 (26.9)

Chronic pulmonary disease	2,240,200 (26.8)	7,652 (25.5)
Hypothyroidism	2,207,232 (26.4)	7,677 (25.6)
Tumor without metastases	2,207,232 (22.9)	4,660 (15.5)
Obesity	1,593,849 (19.0)	4,862 (16.2)
Cerebrovascular disease	1,577,796 (18.8)	5,088 (17.0)

*Values reflect mean (SD) or N (%).

†Average values over Medicare 5% sample cohorts from 2012 to 2021.

‡Values in subset of patients with both electronic health record (EHR) and claims data available for analysis.

§See Supplemental Table 2 for prevalence of all predictors considered.

621 **Table 2. Comparison of Model Statistics***

	Sensitivity	Specificity	PPV	NPV	c-statistic (95% CI)
Medicare 2021 Cohort					
Model 1. Original predictors,† original coefficients	47.9	86.0	9.1	98.5	0.770 (0.767, 0.773)
Model 2. Original predictors,† refit coefficients	52.6	86.1	10.0	98.6	0.795 (0.792, 0.798)
Model 3. New predictors, refit coefficients (LASSO min)	53.8	86.1	10.2	98.7	0.801 (0.799, 0.804)
Model 4. New predictors, refit coefficients (LASSO 1se)	53.2	86.1	10.1	98.6	0.798 (0.795, 0.801)
Model 5. New predictors, refit coefficients, Model 1 predictors required (restricted LASSO min)	53.8	86.1	10.2	98.7	0.801 (0.795, 0.804)
Model 6. New predictors, refit coefficients, Model 1 predictors required (restricted LASSO 1se)	52.9	86.1	10.0	98.6	0.797 (0.794, 0.800)
Mixed Data Cohort					
Optimal Medicare model predictors and coefficients‡					
Predictors in EHR	43.6	85.9	9.2	97.9	0.746 (0.731, 0.762)
Predictors in claims	44.5	86.0	9.4	97.9	0.763 (0.748, 0.777)
Predictors in either EHR or claims	45.7	86.0	9.7	98.0	0.765 (0.750, 0.779)

Optimal Medicare model predictors, refit coefficients					
Predictors in EHR	45.7	86.0	9.7	98.0	0.766 (0.751, 0.780)
Predictors in claims	47.1	86.1	10.0	98.0	0.773 (0.759, 0.787)
Predictors in either EHR or claims	47.6	86.1	10.1	98.0	0.777 (0.762, 0.791)
All potential predictors, refit coefficients					
Predictors in EHR	46.3	86.0	9.8	98.0	0.767 (0.752, 0.781)
Predictors in claims	46.1	86.0	9.8	98.0	0.776 (0.762, 0.790)
Predictors in either EHR or claims	47.0	86.0	10.0	98.0	0.778 (0.764, 0.792)

*Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) calculated using the 85th percentile as the threshold. CI, confidence interval.

†"Original predictors" refers to predictor variables from the original eRADAR model, defined using the available data source. See Supplemental Table 1 definitions.

‡**See Supplemental Table 5 claims-based coefficients.**

Figures.

Figure 1. Accuracy of Claims-Based Algorithms for Predicting Undiagnosed Dementia

C-statistics and 95% confidence intervals are shown from the LASSO min models averaged over the 10 Medicare cohorts overall and stratified by gender and race/ethnicity. The overall c-statistic was 0.801 [0.799, 0.804]. C-statistics were similar for females (0.803 [0.800, 0.807]) and males (0.796 [0.792, 0.801]), but were slightly lower for Asian (0.787 [0.766, 0.807]), Black (0.779 [0.769, 0.790]), Hispanic (0.778 [0.756, 0.800]), and North American Native (0.779 [0.736, 0.822]) than White (0.803 [0.800, 0.806]) individuals.

Figure 2. Sensitivity for Detecting Undiagnosed Dementia at Varying Screening Thresholds

The solid blue line shows the proportion of patients with undiagnosed dementia that would be identified at various risk score screening thresholds using data from the optimal Medicare (LASSO min 2021) model. The dotted blue lines show that 54% of patients with undiagnosed dementia would be correctly identified using 15% as the screening threshold, and 79% of patients with undiagnosed dementia would be correctly identified using 35% as the screening threshold.

Figure 3. Comparison of Model Discrimination Using Different Data Sources

Receiver operating characteristic (ROC) curves are shown under different data source scenarios. Discrimination is slightly lower using electronic health record (EHR) data (blue line) than claims data (red line) with predictors and coefficients from the optimal Medicare model. Discrimination is improved by refitting the optimal Medicare model (green and yellow lines) and by considering new predictors (purple and orange lines) regardless of data source.