



# HHS Public Access

Author manuscript

*J Am Geriatr Soc.* Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

*J Am Geriatr Soc.* 2020 January ; 68(1): 103–111. doi:10.1111/jgs.16182.

## Development and Validation of eRADAR: A Tool using EHR Data to Detect Unrecognized Dementia

Deborah E. Barnes, PhD, MPH<sup>1,2,3</sup>, Jing Zhou, PhD<sup>4</sup>, Rod L. Walker, MS<sup>4</sup>, Eric B. Larson, MD, MPH<sup>5</sup>, Sei J. Lee, MD, MS<sup>3,6</sup>, W. John Boscardin, PhD<sup>2,3,6</sup>, Zachary A. Marcum, PharmD, PhD<sup>7</sup>, Sascha Dublin, MD, PhD.<sup>5,8</sup>

<sup>1</sup>Department of Psychiatry, University of California, San Francisco

<sup>2</sup>Department of Epidemiology & Biostatistics, University of California, San Francisco

<sup>3</sup>San Francisco Veterans Affairs Health Care System

<sup>4</sup>Biostatistics Unit, Kaiser Permanente Washington Health Research Institute

<sup>5</sup>Kaiser Permanente Washington Health Research Institute

<sup>6</sup>Department of Medicine, Division of Geriatrics, University of California, San Francisco

<sup>7</sup>Department of Pharmacy, University of Washington

<sup>8</sup>Department of Epidemiology, University of Washington

### Abstract

**Background:** Early recognition of dementia would allow patients and their families to receive care earlier in the disease process, potentially improving care management and patient outcomes, yet nearly half of patients with dementia are undiagnosed.

**Objective:** To develop and validate an electronic health record (EHR)-based tool to help detect patients with unrecognized dementia (eRADAR).

**Design:** Retrospective cohort study.

**Setting:** Kaiser Permanente Washington (KPWA), an integrated healthcare delivery system.

**Participants:** 16,665 visits among 4,330 participants in the Adult Changes in Thought (ACT) study, who undergo a comprehensive process to detect and diagnose dementia every two years and have linked KPWA EHR data, divided into development (70%) and validation (30%) samples.

---

**Corresponding author:** Deborah E. Barnes, PhD, MPH; 4150 Clement St, MC 151R; San Francisco, CA 94121; Deborah.barnes@ucsf.edu; @DeborahEBarnes.

Authors Contributions:

DEB conceived of the study; contributed to study design and analysis and interpretation of data; and prepared the manuscript.

JZ contributed to study design, conducted data analyses, and provided feedback on the manuscript.

RLW contributed to study design, conducted data analyses, and provided feedback on the manuscript.

EBL acquired subjects and data; contributed to interpretation of data; and provided feedback on the manuscript.

SJL contributed to study conception and design and analysis and interpretation of data and provided feedback on the manuscript.

WJB contributed to data analysis and interpretation and provided feedback on the manuscript.

ZAM contributed to data analysis and interpretation and provided feedback on the manuscript.

SD contributed to study conception and design and analysis and interpretation of data and provided feedback on the manuscript.

**Conflict of Interest:** The authors have no conflicts of interest.

**Measurements:** EHR predictors included demographics, medical diagnoses, vital signs, healthcare utilization and medications within the previous two years. Unrecognized dementia was defined as detection in ACT prior to documentation in the KPWA EHR (i.e., lack of dementia or memory loss diagnosis codes or dementia medication fills).

**Results:** 1,015 ACT visits resulted in a diagnosis of incident dementia, of which 498 (49%) were unrecognized in the KPWA EHR. The final 31-predictor model included markers of dementia-related symptoms (e.g., psychosis diagnoses, anti-depressant fills), healthcare utilization pattern (e.g., emergency department visits); and dementia risk factors (e.g., cerebrovascular disease, diabetes). Discrimination was good in the development (c statistic, 0.78; 95% CI: 0.76, 0.81) and validation (0.81; 0.78, 0.84) samples, and calibration was good based on plots of predicted vs. observed risk. If patients with scores in the top 5% were flagged for additional evaluation, we estimate that 1 in 6 would have dementia.

**Conclusion:** eRADAR uses existing EHR data to detect patients who may have unrecognized dementia with good accuracy.

### Keywords

dementia; decision support techniques; early diagnosis

## INTRODUCTION

There are currently 5.8 million people in the U.S. living with dementia, approximately half of whom have never received a diagnosis.<sup>1,2</sup> Many potential benefits of earlier dementia recognition have been proposed.<sup>3-5</sup> Reversible and treatable causes of cognitive impairment can be identified. Evidence-based collaborative care models can be implemented to enable better management of patient symptoms,<sup>6,7</sup> which could lower risk of emergency department (ED) visits and hospitalizations.<sup>8</sup> Patients may choose to start cholinesterase inhibitors, which may improve symptoms in some patients.<sup>9</sup> Patients can be engaged in planning for their future, and caregivers can be provided with education and support, which can reduce stress and improve well-being.<sup>10,11</sup> Interventions can be implemented to minimize risky behaviors such as driving.<sup>12,13</sup>

There are also potential harms of earlier detection of dementia. These include anxiety/depression; stigma/social isolation; loss of independence such as driving privileges; and negative financial consequences such as denial of long-term care insurance.<sup>3,4</sup>

Given the tension between potential benefits and harms, agencies such as the U.S. Preventive Services Task Force and the U.K. National Institute for Health Care Excellence do not recommend for or against routine screening for dementia.<sup>14,15</sup> However, other organizations such as the Alzheimer's Association and the Gerontological Society of America advocate for early detection and diagnosis so that patients and families can be provided with adequate support and education throughout the disease process.<sup>5,16</sup>

There are many barriers to early dementia diagnosis in the primary care setting, including lack of evidence from clinical trials and lack of demand from patients, caregivers and clinicians.<sup>4,5,17-20</sup> Insufficient time is often identified as the single most important factor by

primary care physicians.<sup>2,18,19,21,22</sup> Several studies also have identified the need for standardized tools and information technology resources to address this barrier and support earlier recognition of Alzheimer's and dementia in primary care.<sup>2,17,22</sup>

We have previously found that patients with unrecognized dementia often have 'warning signs' (i.e., markers of high risk) in the electronic health record (EHR) such as history of stroke, recent ED visits and "no shows" for scheduled clinic visits.<sup>23</sup> The goal of this study was to develop and validate an EHR-based prediction tool to identify patients at high risk of unrecognized dementia, which we call the EHR Risk of Alzheimer's and Dementia Assessment Rule (eRADAR).

## METHODS

### Setting

The setting for this study is Kaiser Permanente Washington (KPWA, formerly Group Health Cooperative). KPWA is a not-for-profit health plan that provides medical coverage and care to more than 710,000 members in Washington State, including nearly 93,000 Medicare Advantage beneficiaries. The two-year retention rate for members age 65 years or older is about 85%. KPWA utilizes the Epic EHR, which was fully deployed at all KPWA clinics and specialty centers as of November 2005. Before this time, healthcare data were recorded and maintained in a variety of other electronic databases, which remain available and have been widely used for research. Diagnosis and procedure data are available dating back to the early 1990s, while electronic prescription data are available back to 1977.

### Study population

Our study population was drawn from the Adult Changes in Thought (ACT) study, an ongoing, prospective cohort study of risk factors for Alzheimer's disease and dementia that is embedded within KPWA.<sup>24</sup> In 1994, ACT enrolled a random sample of Seattle-area KPWA members who were at least 65 years old, community-dwelling and did not have dementia. Additional enrollment waves have been added periodically to maintain the cohort size. ACT study participants are 89% non-Hispanic white and have a median (interquartile range) of 14 (12, 16) years of education.

### Diagnosis of dementia

Every 2 years, ACT participants undergo cognitive screening with the Cognitive Abilities Screening Instrument (CASI).<sup>25</sup> Those with abnormal results undergo in-depth evaluation, including a neuropsychological test battery, physical examination including neurologic assessment, and detailed review of medical records. If no recent imaging results are available, imaging may be ordered. A multidisciplinary consensus committee reviews all available data and assigns diagnoses based on standard research criteria.<sup>26,27</sup>

This study used the ACT visit as the unit of analysis to take advantage of repeated dementia assessments in the ACT cohort. We classified participants' status at each ACT follow-up visit as 1) no dementia, 2) recognized dementia or 3) unrecognized dementia. Specifically, we classified dementia as 'recognized' if the KPWA EHR included any diagnosis codes for

dementia or memory complaints or dementia medication fills during the two years prior to the ACT visit date and as ‘unrecognized’ if none of these criteria were met. We included diagnosis codes for memory complaints in our ‘recognized’ dementia group because our goal was to develop a tool to help clinicians identify patients who were not already recognized as being ‘at risk.’ We excluded baseline visits because individuals with dementia at baseline were not enrolled in ACT. We also excluded visits at which participants were not enrolled in KPWA for the prior two years (because EHR predictors might be incomplete) and visits after KPWA’s transition to International Classification of Diseases 10 (ICD-10) coding (10/1/2015, for data consistency). Our final sample included 16,655 study visits among 4,330 patients.

### Potential EHR predictors of unrecognized dementia

A wide range of potential EHR predictors of unrecognized dementia were considered. Some predictors were informed by our prior work<sup>23</sup> and a priori hypotheses regarding factors that might be associated with increased risk of undiagnosed dementia (such as receiving care for dementia-related symptoms, missing clinic visits and poor medication adherence). Other predictors were based on well-established algorithms for administrative data.<sup>28,29</sup> The period of assessment for most predictors was the two years preceding the ACT visit. Height, weight and blood pressure were not captured in the EHR until 2003; therefore, we used the most recently recorded value from either the EHR or the ACT visit during the preceding three years to minimize missing data. Predictors were grouped into the broad categories of demographics, diagnoses, vital signs, healthcare utilization and medication-related predictors. In addition, to maximize the clinical utility of the tool, we classified predictors as being more or less easily obtainable in most medical systems. A complete list of predictors considered is included in Supplementary Table S1.

#### More easily obtainable EHR predictors

**a) Demographics:** We considered age, gender, and race/ethnicity as potential demographic predictors. Race/ethnicity was dichotomized as non-Hispanic white versus other because the proportion of ACT participants with nonwhite race/ethnicity is small (10%).

**b) Diagnoses.:** We identified comorbid medical conditions using a code list drawn primarily from ICD-9 codes recommended by Elixhauser or Charlson.<sup>28,29</sup> A total of 31 specific conditions were considered including hypertension, myocardial infarction, congestive heart failure, cerebrovascular disease, diabetes, alcohol dependence/abuse, psychoses, depression, traumatic brain injury, tobacco use disorder, atrial fibrillation and gait abnormality.

**c) Vital signs.:** Height, weight and blood pressure were determined based on the most recently recorded EHR value. Height and weight were used to calculate body mass index (BMI, kg/m<sup>2</sup>), which was classified using standard cut-points: underweight (<18.5), normal (18.5–24.9), overweight (25.0–29.9) or obese (≥30). Systolic and diastolic blood pressure (SBP, DBP, mm Hg) were categorized using standard clinical cut-points: low (<90 SBP or <60 DBP), normal (90–139 SBP and 60–89 DBP), high (140–180 SBP or 90–110 DBP) or

very high (>180 SBP or >110 DBP). For <1% of visits, no information on blood pressure or BMI was available in the preceding three years. In these cases, we imputed values for the missing data using fully conditional specification methods available in SAS, version 9.4 (SAS Institute, Inc., Cary, NC).

**d) Healthcare utilization.** We classified the number and type of interactions with the healthcare system including outpatient, ED and urgent care visits, and hospitalizations. In addition, we identified visits and hospitalizations for accidents/injuries (which could reflect dementia-related challenges with daily activities) and we determined whether patients had received specialty services that might be related to symptoms of dementia such as social work; home health; physical therapy; mental health; speech, language and learning evaluation (the primary route for cognitive or neuropsychologic testing within KPWA); or neurology.

**e) Medication-related predictors.** We used electronic pharmacy dispensing data to examine fills for medications that could be related to dementia symptoms including tricyclic and non-tricyclic (TCA) anti-depressants and sedative/hypnotic medications.

#### **Less easily obtainable EHR predictors**

**a) Healthcare utilization.** We created a separate variable for hospitalizations/ED visits related to ambulatory care sensitive (ACS) conditions (which are often considered a marker of potentially avoidable admissions).<sup>30</sup> In addition, we drew on KPWA scheduling and clinical databases to identify “no shows” for scheduled clinic visits.

**b) Medication-related predictors.** We examined medication adherence for three categories of commonly prescribed medications (oral hypoglycemic agents, antihypertensive medications and statins). Non-adherence was defined as <80 percent of days covered (PDC), consistent with national quality measures.<sup>31</sup> Individuals who did not take these medications or had <80% PDC were classified as adherent. Four variables were created to reflect non-adherence for each class separately as well as overall.

#### **Analytic approach**

Logistic regression with LASSO (least absolute shrinkage and selection operator) penalty<sup>32</sup> was used to identify EHR predictors of unrecognized dementia versus no dementia and to build a prediction model to estimate the probability of unrecognized dementia for each person at each visit. Visits classified as ‘recognized dementia’ were excluded since our goal was to identify patients with unrecognized dementia to target for further evaluation.

The LASSO penalization factor selects predictors by shrinking coefficients for weaker predictors toward zero, excluding predictors with estimated zero coefficients from the final prediction model. Simulation studies suggest that LASSO leads to less overfitting and more accurate prediction models than more traditional methods such as stepwise selection.<sup>33</sup>

For model development and validation, we randomly divided the data into training (70%) and test (30%) samples. To minimize overfitting, we used the training sample to select the LASSO tuning parameter  $\lambda$  that minimized the binomial deviance via 10-fold cross

validation. We generated the final model by fitting the logistic regression with LASSO to the entire training sample using the selected tuning parameter. We implemented the above procedure using the glmnet package (version 2.0.16) in R (version 3.4.4). We assessed calibration and discrimination performance in the training and test samples graphically<sup>34</sup> and through computation of c-statistics (area under the receiver operating characteristic [ROC] curve [AUC] for a binary outcome). Confidence intervals for AUC estimates were calculated via bootstrap with 10,000 replications.

Two models were created. We first considered the full set of 64 predictors described in the preceding section. Next, to maximize the potential portability of the model into other healthcare systems, we rebuilt the model after excluding predictors likely to be difficult to obtain in some settings (specifically, variables based on ACS conditions, no shows and medication non-adherence).

For sensitivity analyses, we explored alternative modeling using other machine learning methods such as ridge regression, decision trees, random forests, gradient boosting, support vector machines, and neural networks.<sup>35</sup> We also estimated another prediction model using all dementia cases (both recognized and unrecognized) and compared its performance with that of our primary model.

### **Role of the funding source**

Funders were not involved with the design, conduct or reporting of results.

### **IRB approval**

All study procedures were approved by Institutional Review Boards at the University of California, San Francisco and Kaiser Permanente Washington and by the San Francisco Veterans Affairs Research & Development Committee.

## **RESULTS**

Our final sample included 16,655 ACT visits among 4,330 patients. At the patient level, 23.4% developed dementia, 24.7% died, 7.1% withdrew from ACT, 34.6% reached the end of our study period without developing dementia, 4.7% disenrolled from KPWA, and 5.5% were lost to follow-up (e.g., did not return for their biennial visits). At the visit level, 1,015 visits resulted in an ACT dementia diagnosis, of which about half (n=498) were unrecognized in the EHR (overall, 49%; 1996–2000, 54%; 2001–2005: 59%; 2006–2010: 46%; 2011–2015: 38%). The mean age of participants over the visits was 80 years, and 60% of visits were in female participants. Table 1 shows the prevalence of selected predictors overall and stratified by dementia status at the time of the visit. Data for all predictors considered are included in Supplementary Table S1.

The final coefficients for the restricted and full models are shown in Table 2. The restricted model included a combination of demographics (older age, male sex); medical diagnoses (e.g., psychoses, diabetes, congestive heart failure, cerebrovascular disease, gait abnormalities); healthcare utilization (e.g., emergency department or speech therapy visits); vital signs (e.g., underweight); and medication-related predictors (e.g., antidepressant fills



other than TCAs) (Figure 1). The restricted model c-statistic was 0.78 (95% CI: 0.76, 0.81) in the training sample and 0.81 (0.78, 0.84) in the test sample. The c-statistic in the test sample was consistent over time (1996–2000: 0.83 [0.75, 0.90]; 2001–2005: 0.78 [0.73, 0.83]; 2006–2010: 0.79 [0.72, 0.85]; 2011–2015: 0.83 [0.76, 0.90]).

In the full model, most predictors and coefficient values were similar to those for the restricted model (Table 2). The key differences were that some variables were no longer included (diabetes, lymphoma, or depression diagnoses; home health visits; sedative/hypnotic medication fills) while others were added (clinic no shows; ACS hospitalizations and ED visits; non-adherence to hypoglycemic, antihypertensive or statin medications). The full model c-statistic was 0.79 (0.65, 0.81) in the training sample and 0.81 (0.77, 0.84) in the test sample.

Figure 2 shows the performance characteristics of the final restricted model in the training and test samples (see Supplementary Figure S1 for full model). The ROC curves and AUC estimates (Figure 2a) suggest good discrimination between patients with no dementia and unrecognized dementia. The calibration plots (Figure 2b) suggest reasonable correspondence between observed and predicted risk across the full range of scores. Overall, there is no evidence of overfitting based on a comparison of discrimination and calibration between training and test samples.

Table 3 shows performance characteristics of the final restricted model when using different predicted risk cut-offs for classifying subjects as having unrecognized dementia at a given visit (see Supplementary Table S3 for full model). For example, using a 95<sup>th</sup> percentile cut-off in the training sample yielded a sensitivity of 22%, specificity of 96%, positive predictive value (PPV) of 16% and negative predictive value (NPV) of 98% in the test sample. This suggests that, if the top 5% of patients were evaluated, 16% (roughly 1 in 6 patients) would be predicted to have unrecognized dementia. In contrast, the overall visit-level prevalence of unrecognized dementia was 3%.

Exploration of other machine learning methods did not produce better performing models compared to our primary approach (c-statistics from restricted model test sample: ridge regression: 0.81; random forest: 0.76; boosting: 0.81; SVM: 0.67; neural network: 0.64). Using all dementia cases (rather than restricting to unrecognized dementia) also did not improve performance (restricted model: 0.80; full model: 0.80).

## DISCUSSION

In this study of more than 16,000 visits between 1995 and 2015 among 4,330 older adults from the ACT study,<sup>24</sup> we found that about half of participants who were diagnosed with incident dementia through ACT appeared to be unrecognized by the healthcare system. The proportion who were unrecognized declined slightly over time but was high in all study years. These findings are consistent with prior studies<sup>2,5,17,36</sup> and suggest that under-recognition of dementia remains a major concern.

We also found that information that is readily available in the EHR can be used to detect patients who may have unrecognized dementia with good accuracy. Key predictors included

patterns of health care utilization, dementia-related symptoms, and dementia risk factors. We also found that variables such as poor medication adherence and clinic visit ‘no shows’ were associated with increased risk of unrecognized dementia, although adding these more novel predictors (which may be difficult to extract in some EHRs) did not improve prognostic accuracy compared to a model that only considered more easily extractable predictors. This may be explained by correlations between predictors that enabled substitutions without loss of prognostic performance. Model accuracy was slightly better in the test sample, suggesting that over-fitting is not a concern.

We also examined the trade-offs between sensitivity and PPV to understand the potential impact of implementing eRADAR in clinical practice. If the 85<sup>th</sup> percentile were set as a threshold for recommending follow-up evaluation (for example, a phone or office visit to assess memory), sensitivity would be 47% and PPV 10%—that is, we would detect nearly half of patients with undiagnosed dementia, and about 1 in 10 people evaluated would have dementia. For context, 12% of screening mammograms are interpreted as abnormal, with 87% sensitivity and 4% PPV (~23 people evaluated with diagnostic mammograms and/or biopsies to detect 1 breast cancer).<sup>37</sup> For fecal immunochemical testing (FIT), a recent community-based study found 74% sensitivity and 2% PPV using a 20 mcg/g cut-off (~52 people evaluated via colonoscopy to detect 1 case of colorectal cancer).<sup>38</sup> Thus, eRADAR’s performance is comparable to other common clinical procedures.

To our knowledge, this is the first study to develop a risk prediction model that uses readily available EHR data to detect patients with unrecognized dementia. Most existing models focus on predicting future risk of developing dementia.<sup>39,40</sup> For example, our Dementia Screening Indicator<sup>41</sup> is a simple tool designed to identify older adults with an elevated risk of developing dementia within 6 years. These tools may be optimal for identifying patients to target for risk reduction interventions, rather than detecting patients with unrecognized dementia, since they focus on predicting future risk.

Several other tools have recently been developed that use text-based EHR data to more accurately differentiate between patients with and without dementia.<sup>42–44</sup> For example, Amra et al.<sup>42</sup> developed an algorithm that uses EHR text words (such as “cognitive impairment,” “impaired memory” or “difficulty concentrating”) to discriminate between patients with and without dementia. A limitation is that many clinical settings may not have the resources or technical capabilities to extract information from free text fields such as clinic notes. Also, such approaches are likely to identify patients who are already known by clinicians to potentially have dementia. In contrast, our tool is designed to identify patients with dementia who have not yet been recognized by their clinicians.

Another limitation of prior studies is that most models were developed to predict diagnosed dementia and have assumed that results would apply to undiagnosed dementia.<sup>43–45</sup> However, our prior work suggests that patients with undiagnosed dementia often have values that fall in between patients with diagnosed dementia and those without dementia.<sup>23</sup> Thus, models that are optimized to predict diagnosed dementia may be less accurate for identifying unrecognized dementia cases.



This study has several strengths. We examined a large sample that included 16,655 visits among 4,330 patients and nearly 500 cases of unrecognized dementia. In addition, we examined a wide range of potential EHR predictors based on a conceptual model developed a priori.

We also acknowledge several limitations. 1) Our definition of unrecognized dementia relied on EHR diagnosis codes and medication fills. It is possible that some clinicians were aware of a patient's dementia status but did not code or prescribe medications for it, although our pilot work confirmed that 90% of cases were unrecognized based on detailed chart review.<sup>23</sup> 2) Generalizability may be limited given our sample, which was primarily white, well-educated, English speaking, and from one healthcare system in one region of the country. Patterns of healthcare utilization, access to services and medication compliance may differ in those with lower socioeconomic status or education levels, and results may not be readily transported outside the U.S. where practices for screening, diagnosis and coding of dementia may differ. 3) If implemented, the model would not identify all patients with undiagnosed dementia, and those not identified would not be targeted for further evaluation. 4) Some predictors in the final model do not make intuitive sense; for example, solid tumor without metastasis is associated with a lower risk of undiagnosed dementia. These patients may be more likely to receive a diagnosis due to greater interaction with the health care system. We also note that prognostic models can be accurate even when counterintuitive. 5) There are other predictors that we did not consider in this model, such as antipsychotic and anticholinergic medications, non-compliance with other types of medications, laboratory test results, and hospital visits with delirium or intensive care unit stays. We also did not use techniques such as natural language processing to examine clinical note fields. Future studies should determine whether the accuracy of our model is improved by including additional predictors or using alternative techniques. It also would be of great value to validate our model in new databases or through a prospective validation study within a real-world population.

We recognize that there are many barriers to dementia diagnosis,<sup>5,17,19,21,46</sup> and that the idea of applying a tool such as eRADAR to detect cases of unrecognized dementia may raise concerns among some patients or clinicians. For example, one study found that nearly half of patients with positive dementia screens declined further evaluation.<sup>47</sup> Implementing an EHR-based tool will require careful design and should incorporate input and guidance from patients, caregivers and clinicians to address barriers and minimize the potential for unintended adverse consequences. It is likely that both clinicians and patients will need enhanced support and resources if more cases of dementia are detected.

In summary, we have developed and internally validated eRADAR, a tool that uses readily available EHR data to identify patients who may have unrecognized dementia with good accuracy. Future studies should explore the optimal approach to implementing eRADAR, which could involve applying it at the point of care (e.g. an EHR-based alert that could fire during a clinic visit) or providing risk score information to clinical teams to support proactive outreach to patients outside of scheduled visits.<sup>48</sup> Future studies should assess not only benefits but also potential downstream costs or burdens to the patient, family and healthcare system.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

**Sponsor's Role:** Sponsors were not involved with the design, methods, subject recruitment, data collection, analysis or preparation of this manuscript.

Preliminary results presented at the Gerontological Society of America meeting in November 2018 and the Alzheimer's Association International Conference in July 2019. Funding provided by the National Institutes of Health (R56 AG056417; R24 AG045050) and Tideswell at UCSF. Data used in this study were collected as part of the Adult Changes in Thought study (U01 AG006781). ZAM was supported by an Agency for Healthcare Research & Quality grant (K12HS022982).

## REFERENCES

1. Alzheimer's Association. 2019 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2019;15(3):321–387.
2. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. *Alzheimer Dis Assoc Disord*. 2009;23(4):306–314. [PubMed: 19568149]
3. Borson S, Frank L, Bayley PJ, et al. Improving dementia care: the role of screening and detection of cognitive impairment. *Alzheimers Dement*. 2013;9(2):151–159. [PubMed: 23375564]
4. Dubois B, Padovani A, Scheltens P, Rossi A, Dell'Agnello G. Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. *J Alzheimers Dis*. 2016;49(3):617–631. [PubMed: 26484931]
5. The Gerontological Society of America. The Gerontological Society of America Workgroup on Cognitive Impairment Detection and Earlier Diagnosis: Report and Recommendations. <https://www.geron.org/images/gsa/documents/gsaciworkgroup2015report.pdf>. Published 2015 Accessed.
6. 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. [PubMed: 16684985]
7. 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. [PubMed: 28364258]
8. 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. [PubMed: 23305822]
9. Birks JS, Harvey RJ. Donepezil for dementia due to Alzheimer's disease. *Cochrane Database Syst Rev*. 2018;6:CD001190.
10. Health Quality Ontario. Caregiver- and patient-directed interventions for dementia: an evidence-based analysis. *Ont Health Technol Assess Ser*. 2008;8(4):1–98.
11. 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. [PubMed: 21631819]
12. 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. [PubMed: 27253366]
13. 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. [PubMed: 27171906]
14. Lin JS, O'Connor E, Rossom R, et al. Screening for cognitive impairment in older adults: An evidence update for the U.S. Preventive Services Task Force. Evidence Report No. 107 AHRQ Publication No. 14–05198-EF-1. Rockville, MD: Agency for Healthcare Research and Quality;2013.

15. National Institute for Health Care Excellence. Dementia: assessment, management and support for people living with dementia and their carers. [nice.org.uk/guidance/ng97](https://www.nice.org.uk/guidance/ng97). Published 2018 Accessed 5/10/2019.
16. Cordell CB, Borson S, Boustani M, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimers Dement*. 2013;9(2):141–150. [PubMed: 23265826]
17. Aminzadeh F, Molnar FJ, Dalziel WB, Ayotte D. A review of barriers and enablers to diagnosis and management of persons with dementia in primary care. *Can Geriatr J*. 2012;15(3):85–94. [PubMed: 23259021]
18. 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. *Gerontologist*. 2004;44(4):500–507. [PubMed: 15331807]
19. 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. [PubMed: 17823840]
20. Fowler NR, Perkins AJ, Turchan HA, et al. Older primary care patients' attitudes and willingness to screen for dementia. *J Aging Res*. 2015;2015:423265.
21. Hansen EC, Hughes C, Routley G, Robinson AL. General practitioners' experiences and understandings of diagnosing dementia: factors impacting on early diagnosis. *Soc Sci Med*. 2008;67(11):1776–1783. [PubMed: 18945531]
22. Iliffe S, Manthorpe J, Eden A. Sooner or later? Issues in the early diagnosis of dementia in general practice: a qualitative study. *Fam Pract*. 2003;20(4):376–381. [PubMed: 12876106]
23. Lee SJ, Larson EB, Dublin S, Walker R, Marcum Z, Barnes D. A Cohort Study of Healthcare Utilization in Older Adults with Undiagnosed Dementia. *J Gen Intern Med*. 2018;33(1):13–15. [PubMed: 28849435]
24. Kukull WA, Higdon R, Bowen JD, et al. Dementia and Alzheimer disease incidence: a prospective cohort study. *Arch Neurol*. 2002;59(11):1737–1746. [PubMed: 12433261]
25. Teng EL, Hasegawa K, Homma A, et al. The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. *Int Psychogeriatr*. 1994;6(1):45–58; discussion 62. [PubMed: 8054493]
26. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed*. Washington, DC: American Psychiatric Association;1994.
27. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939–944. [PubMed: 6610841]
28. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–383. [PubMed: 3558716]
29. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8–27. [PubMed: 9431328]
30. Sanderson C, Dixon J. Conditions for which onset or hospital admission is potentially preventable by timely and effective ambulatory care. *J Health Serv Res Policy*. 2000;5(4):222–230. [PubMed: 11184959]
31. Marcum ZA, Sevick MA, Handler SM. Medication nonadherence: a diagnosable and treatable medical condition. *JAMA*. 2013;309(20):2105–2106. [PubMed: 23695479]
32. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological)*. 1996;58(1):267–288.
33. Steyerberg EW, Eijkemans MJ, Harrell FE Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med*. 2000;19(8):1059–1079. [PubMed: 10790680]
34. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128–138. [PubMed: 20010215]

35. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*, 2nd ed. New York: Springer; 2009.
36. Eichler T, Thyrian JR, Hertel J, et al. Rates of formal diagnosis in people screened positive for dementia in primary care: results of the DelpHi-Trial. *J Alzheimers Dis*. 2014;42(2):451–458. [PubMed: 24898640]
37. Lehman CD, Arao RF, Sprague BL, et al. National Performance Benchmarks for Modern Screening Digital Mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology*. 2017;283(1):49–58. [PubMed: 27918707]
38. Selby K, Jensen CD, Lee JK, et al. Influence of Varying Quantitative Fecal Immunochemical Test Positivity Thresholds on Colorectal Cancer Detection: A Community-Based Cohort Study. *Ann Intern Med*. 2018;169(7):439–447. [PubMed: 30242328]
39. Hou XH, Feng L, Zhang C, Cao XP, Tan L, Yu JT. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry*. 2018.
40. Tang EY, Harrison SL, Errington L, et al. Current Developments in Dementia Risk Prediction Modelling: An Updated Systematic Review. *PLoS One*. 2015;10(9):e0136181.
41. Barnes DE, Beiser AS, Lee A, et al. Development and validation of a brief dementia screening indicator for primary care. *Alzheimers Dement*. 2014;10(6):656–665 e651. [PubMed: 24491321]
42. Amra S, O'Horo JC, Singh TD, et al. Derivation and validation of the automated search algorithms to identify cognitive impairment and dementia in electronic health records. *J Crit Care*. 2017;37:202–205. [PubMed: 27969571]
43. Jammeh EA, Carroll CB, Pearson SW, et al. Machine-learning based identification of undiagnosed dementia in primary care: a feasibility study. *BJGP Open* 2018;2(2):1–13.
44. Reuben DB, Hackbarth AS, Wenger NS, Tan ZS, Jennings LA. An Automated Approach to Identifying Patients with Dementia Using Electronic Medical Records. *J Am Geriatr Soc*. 2017;65(3):658–659. [PubMed: 28152164]
45. Wray LO, Wade M, Beehler GP, Hershey LA, Vair CL. A program to improve detection of undiagnosed dementia in primary care and its association with healthcare utilization. *Am J Geriatr Psychiatry*. 2014;22(11):1282–1291. [PubMed: 23954037]
46. Ahmad S, Orrell M, Iliffe S, Gracie A. GPs' attitudes, awareness, and practice regarding early diagnosis of dementia. *Br J Gen Pract*. 2010;60(578):e360–365. [PubMed: 20849686]
47. Boustani M, Perkins AJ, Fox C, et al. Who refuses the diagnostic assessment for dementia in primary care? *Int J Geriatr Psychiatry*. 2006;21(6):556–563. [PubMed: 16783796]
48. Fowler NR, Morrow L, Chiappetta L, et al. Cognitive testing in older primary care patients: A cluster-randomized trial. *Alzheimers Dement (Amst)*. 2015;1(3):349–357. [PubMed: 26380844]

**Dementia Risk Factors**

- Congestive heart failure
- Stroke
- Diabetes
- Traumatic brain injury
- Gait abnormality

**Dementia-Related Symptoms**

- Psychosis
- Non-tricyclic anti-depressant fills
- Underweight

**Healthcare Utilization Patterns**

- Emergency department visits
- Clinic 'no shows
- Potentially avoidable hospitalizations
- Medication non-adherence

### **Dementia Risk Factors**

- Congestive heart failure
- Stroke
- Diabetes
- Traumatic brain injury
- Gait abnormality

### **Dementia-Related Symptoms**

- Psychosis
- Non-tricyclic anti-depressant fills
- Underweight

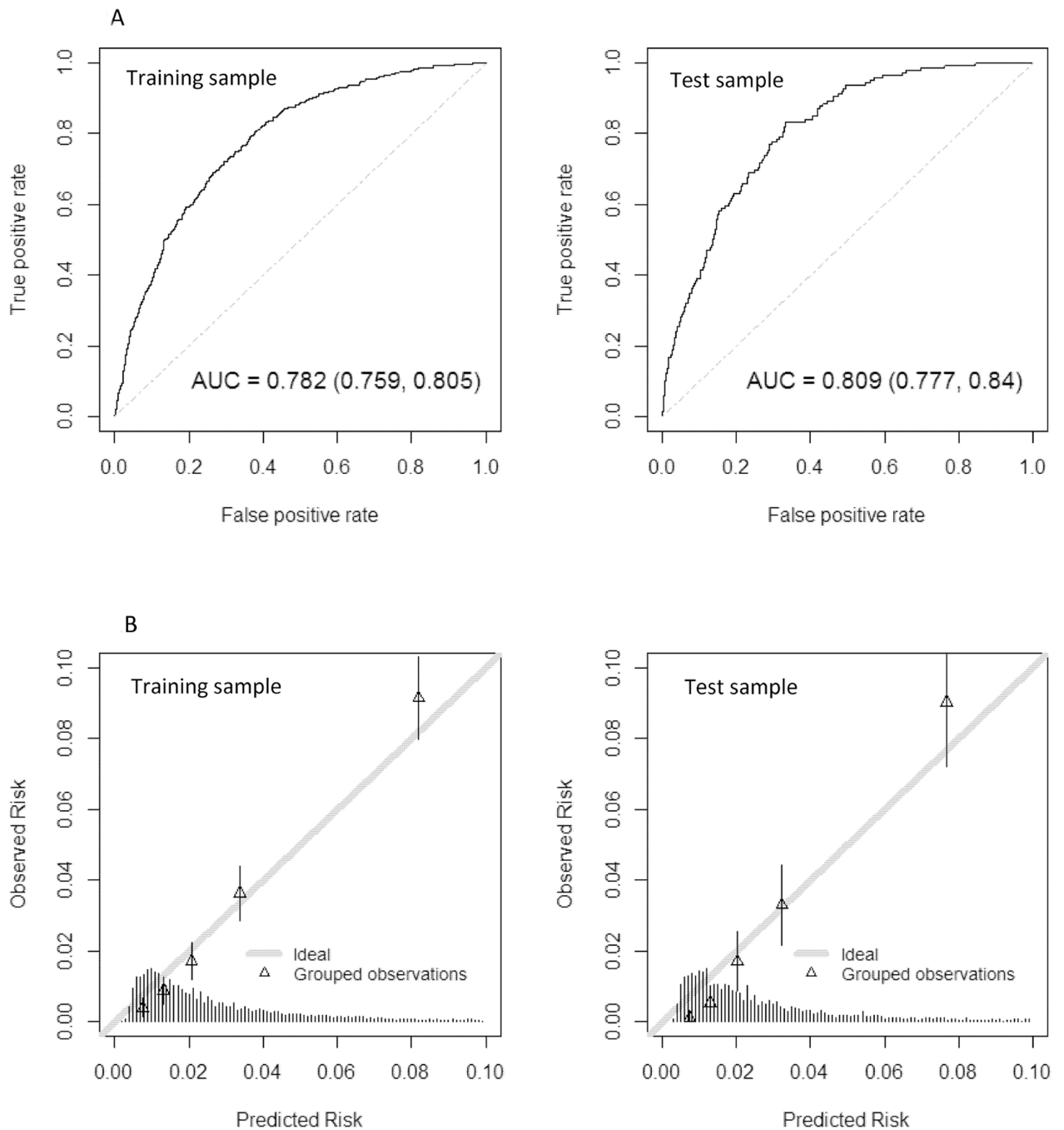
### **Healthcare Utilization Patterns**

- Emergency department visits
- Clinic 'no shows
- Potentially avoidable hospitalizations
- Medication non-adherence

**Figure 1.**

Key predictors of undiagnosed dementia in the electronic health record Risk of Alzheimer's and Dementia Assessment Rule (eRADAR) included dementia risk factors, dementia-related symptoms, and healthcare utilization patterns.





**Figure 2.**

The performance characteristics of the final restricted eRADAR model are shown for the training sample (left panels) and test sample (right panels). Figure 2a shows the receiver operating characteristic (ROC) curves with c-statistics. The ROC curve plots the sensitivity (true positive rate) against 1 - specificity (false positive rate) for consecutive cutoffs for the probability of unrecognized dementia. The c-statistic reflects the area under the ROC curve (AUC), with values of 0.5 reflecting prediction no better than chance and 1 reflecting perfect prediction.

Figure 2b provides a graphical assessment of calibration (i.e., the extent to which predicted risk matches actual risk). The mean predicted probability of unrecognized dementia is plotted against the observed proportion of unrecognized dementia cases by quintiles of predicted risk. The figure shows close alignment between observed and predicted values (ideal calibration aligns with the 45-degree line). The histogram at the bottom shows the distribution of the predicted risks. Most visits had relatively low predicted risk of undiagnosed dementia (<5%). Supplementary Figure S1 provides similar figures for the model that considered the full set of predictors.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1.**

Prevalence of Electronic Health Record (EHR) Predictors Included in Final Models by Visit Status

EHR Predictors	Visit Status*			
	Overall (n=16,655)	No dementia (n=15,640)	Unrecognized dementia (n=498)	Recognized dementia (n=517)
<b>Demographics</b>				
Age in years, mean (SD)	80.1 (6.6)	79.7 (6.5)	85.4 (6.0)	85.0 (6.2)
Female	10,044 (60)	9,417 (60)	304 (61)	323 (62)
<b>Medical diagnoses, past 2 years</b>				
Congestive heart failure	2,127 (13)	1,840 (12)	138 (28)	149 (29)
Cerebrovascular disease	2,028 (12)	1,726 (11)	107 (21)	195 (38)
Diabetes, any	2,425 (15)	2,243 (14)	99 (20)	83 (16)
Diabetes, complex	1,205 (7)	1,090 (7)	63 (13)	52 (10)
Chronic pulmonary disease	3,140 (19)	2,919 (19)	94 (19)	127 (25)
Hypothyroidism	2,237 (13)	2,026 (13)	95 (19)	116 (22)
Renal failure	1,761 (11)	1,587 (10)	67 (13)	107 (21)
Lymphoma	183 (1)	167 (1)	7 (1)	9 (2)
Solid tumor w/o metastases	3,894 (23)	3,656 (23)	99 (20)	139 (27)
Rheumatoid arthritis	953 (6)	877 (6)	26 (5)	50 (10)
Weight loss	116 (1)	91 (1)	3 (1)	22 (4)
Fluid and electrolyte disorders	2,215 (13)	1,914 (12)	99 (20)	202 (39)
Blood loss anemia	737 (4)	661 (4)	30 (6)	46 (9)
Psychoses	471 (3)	308 (2)	35 (7)	128 (24)
Depression	2,573 (15)	2,235 (14)	129 (26)	209 (40)
Traumatic brain injury	252 (2)	211 (1)	22 (4)	19 (4)
Tobacco use disorder	1,213 (7)	1,132 (7)	24 (5)	57 (11)
Atrial fibrillation	2,513 (15)	2,260 (14)	117 (23)	136 (26)
Gait abnormality	1,534 (9)	1,292 (8)	87 (17)	155 (30)
<b>Vital signs, most recent, past 3 years</b>				
Underweight (BMI<18.5)	289 (2)	243 (2)	20 (4)	26 (5)
Obese (BMI ≥ 30)	3,625 (22)	3,478 (22)	72 (15)	75 (15)
High BP (≥ 140 SBP or 90 DBP)	457 (3)	430 (3)	13 (3)	14 (3)
<b>Healthcare utilization, past 2 years</b>				
1 outpatient visit	16,453 (99)	15,449 (99)	490 (98)	260 (99)
1 emergency department visit	4,606 (28)	4,095 (26)	230 (46)	281 (54)
1 home health visit	1,617 (10)	1,380 (9)	89 (18)	148 (29)
1 physical therapy visit	6,293 (38)	5,902 (38)	181 (36)	210 (41)
1 speech, language and learning visit	478 (3)	343 (2)	24 (5)	111 (21)
<b>Medication-related predictors, past 2 years</b>				
Fill for non-TCA antidepressant	2,426 (15)	2,102 (12)	135 (26)	189 (35)
Fill for sedative/hypnotic	4,472 (27)	4,144 (21)	150 (22)	178 (28)

EHR Predictors	Visit Status*			
	Overall (n=16,655)	No dementia (n=15,640)	Unrecognized dementia (n=498)	Recognized dementia (n=517)
<b>Additional predictors considered for full model</b>				
1 clinic "no shows"	5,143 (31)	4,666 (30)	210 (42)	267 (52)
1 ACS hospitalizations	166 (1)	138 (1)	7 (1)	21 (4)
1 ACS emergency visits	1,109 (7)	999 (6)	50 (10)	60 (12)
Medication non-adherence <sup>a</sup>	2,928 (18)	2,662 (17)	131 (26)	135 (26)

\* Values are number (%) or mean (standard deviation). For <1% of visits, no information on blood pressure or BMI was available in the preceding three years. In these cases, we imputed values for the missing data using fully conditional specification methods available in SAS, version 9.4 (SAS Institute, Inc., Cary, NC). ACS, ambulatory care sensitive; BMI, body mass index; DBP, diastolic blood pressure; TCA, tricyclic antidepressant; SBP, systolic blood pressure; SD, standard deviation. Table includes predictors retained in final models. See Supplementary Table S1 for prevalence of all predictors considered.

<sup>a</sup> defined as <80 percent of days covered (PDC) for oral anti-diabetics, anti-hypertensives or statins

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.**

Final Restricted and Full eRADAR Model Coefficients for Predicting Undiagnosed Dementia

Predictor	Restricted Model	Full Model
Intercept	-11.83	-11.86
<b>Demographics</b>		
Age (per year)	0.11	0.11
Female	-0.10	-0.07
<b>Diagnoses, past 2 years</b>		
Congestive heart failure	0.28	0.23
Cerebrovascular disease	0.18	0.16
Diabetes, any	0.06	
Diabetes, complex	0.34	0.29
Chronic pulmonary disease	-0.11	-0.10
Hypothyroidism	0.05	0.02
Renal failure	-0.15	-0.13
Lymphoma	0.01	
Solid tumor without metastases	-0.17	-0.16
Rheumatoid arthritis	-0.05	-0.01
Weight loss	-0.34	-0.31
Fluid and electrolyte disorders	-0.07	-0.06
Blood loss anemia	-0.25	-0.22
Psychoses	0.46	0.43
Depression	0.03	
Traumatic brain injury	0.20	0.17
Tobacco use disorder	-0.15	-0.12
Atrial fibrillation	0.03	0.02
Gait abnormality	0.24	0.20
<b>Vital signs, most recent, past 3 years</b>		
Underweight (BMI < 18.5)	0.29	0.26
Obese (BMI ≥ 30)	-0.11	-0.09
High blood pressure (≥ 140 SBP or ≥ 90 DBP)	-0.09	-0.07
<b>Healthcare utilization, past 2 years</b>		
1 outpatient visit	-0.55	-0.53
1 emergency department visit	0.34	0.30
1 home health visit	0.00	
1 physical therapy visit	-0.13	-0.14
1 speech, language and learning visit	0.47	0.42
<b>Medications, past 2 years</b>		
Fill for non-tricyclic antidepressant	0.58	0.55
Fill for sedative-hypnotic	0.03	
<b>Additional predictors considered for full model</b>		
1 clinic "no show"		0.18

Predictor	Restricted Model	Full Model
1 ACS hospitalizations		0.10
1 ACS emergency visits		-0.06
Medication non-adherence <sup>a</sup>		0.29

Values reflect truncated coefficients from the final restricted and full models. Full coefficients are included in Supplementary Table S2. ACS, ambulatory care sensitive; BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

<sup>a</sup> defined as <80% of days covered (PDC) for oral anti-diabetics, anti-hypertensives or statins

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 3.**

Performance of Restricted Model Using Different Decision Rules

Risk cut-off percentile*	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
99 <sup>th</sup>	6.5	99.4	23.7	97.2
95 <sup>th</sup>	22.5	96.4	15.7	97.6
90 <sup>th</sup>	36.2	91.6	11.5	97.9
85 <sup>th</sup>	47.1	87.2	10.0	98.2
80 <sup>th</sup>	59.4	82.9	9.5	98.5
75 <sup>th</sup>	65.9	78.0	8.3	98.7
70 <sup>th</sup>	72.5	72.9	7.5	98.9
65 <sup>th</sup>	79.0	67.5	6.8	99.1
60 <sup>th</sup>	83.3	62.6	6.3	99.2
55 <sup>th</sup>	87.7	56.9	5.8	99.4
50 <sup>th</sup>	91.3	51.9	5.4	99.5

NPV, negative predictive value; PPV, positive predictive value

\* Percentile of the predicted risk distribution observed in the training sample. Performance characteristics are then shown in the test sample when using the given risk cut-off for classifying as unrecognized dementia. Values for the model that considered the full set of predictors are included in Supplementary Table S3.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript