Leveraging electronic health records to examine differential clinical outcomes in people with Alzheimer's Disease

Shruthi Venkatesh, BS[1] *, Linshanshan Wang, MS[2,3] *, Michele Morris, BS[4], Mohammed Moro, MS[2], Ratnam Srivastava, MS[2], Yunqing Han, MS[5], Riddhi Patira, MD[1], Sarah Berman, MD, PhD[1], Oscar Lopez, MD[1], Shyam Visweswaran, MD, PhD[4], Tianrun Cai, MD[5], Tianxi Cai, ScD[2,3] #, Zongqi Xia, MD, PhD[1] #

1. Department of Neurology, University of Pittsburgh, Pittsburgh, PA
2. Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA
3. Department of Biomedical Informatics, Harvard Medical School, Boston, MA
4. Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA
5. Department of Rheumatology, Brigham and Women's Hospital, Boston, MA

*** Share co-first authorship**

 # Share co-senior authorship

Corresponding Authors:

Zongqi Xia zxia1@post.harvard.edu

Tianxi Cai tcai@hsph.harvard.edu

ABSTRACT

BACKGROUND. Alzheimer's disease (AD) carries a high societal burden inequitably distributed across demographic groups.

OBJECTIVE. To examine differences in readily ascertainable clinical outcomes of AD decline among demographic groups.

METHODS. Leveraging the electronic health record (EHR) data (1994-2022) of patients with ≥1 diagnosis code for AD or related dementia from two large healthcare systems, we applied a knowledge graph-guided unsupervised phenotyping algorithm to predict AD diagnosis status and validated using gold-standard chart-reviewed and registry-derived diagnosis labels. After excluding patients with <24 months of data or who were admitted to nursing homes prior to AD diagnosis, we performed survival analyses at each healthcare system to assess the time to two readily ascertainable clinical outcomes of AD decline (*i.e.,* nursing home admission, death), stratified by demographic groups and accounting for baseline covariates (*e.g.,* age, gender, race, ethnicity, healthcare utilization, and comorbidities). We then performed a fixed-effects meta-analysis of the survival analysis data from both healthcare systems.

RESULTS. The AD diagnosis phenotyping algorithm demonstrated high accuracy in identifying AD patients across both healthcare systems (AUROC score range: 0.835-0.923). Of the 34,181 AD patients in both healthcare systems (62% women, 90% non-Hispanic White, 80.39±9.28 years of age at AD diagnosis), 32% were admitted to nursing homes and 50% died during follow-up. In the fixed-effect meta-analysis, non-Hispanic White patients had a lower risk of nursing home admission (HR[95% CI]=0.825[0.776-0.877], $p<0.001$) and higher risk of death (HR[95% CI]=1.381[1.283-1.487], $p<.0001$) than racial and ethnic minorities. There was no difference between women and men in their risk of nursing home admission (HR[95% CI]=1.008[0.967-

1.050], *p*=.762), but women had a lower risk of death (HR[95% CI]=0.873[0.837-0.910], *p*<.0001) than men.

CONCLUSION. This study creates two large EHR-based AD cohorts and adds to the real-world evidence of demographic differences in clinical AD decline, which could potentially inform individual clinical management and future public health policies.

INTRODUCTION

Late-onset Alzheimer's disease (AD) is the leading cause of dementia and neurological disability affecting ~6.9 million people in the United States.[1–3] People with AD experience variable rates of decline with differences among demographic groups. The likelihood of nursing home admission, indicating functional independence loss, is higher among racial and ethnic minorities with AD.[4,5] In contrast, post-diagnosis survival is shorter for men and non-Hispanic White (NHW) individuals.[6–10] While sociodemographic factors and comorbidities influence differential AD decline, there remains a gap in identifying potentially modifiable factors and quantifying the extent to which they drive AD decline.[9,11–16]

While traditional epidemiological studies laid the population-health foundation in examining differential AD decline, large-scale real-world clinical data could enhance generalizability. With growing adoption and standardization, electronic health record (EHR) data contain the longitudinal clinical profiles of demographically broad patient populations and complement traditional data-sources for large-scale clinical investigations. However, accurate identification of target populations (*e.g.,* AD) remains a challenge, limiting the effective use of EHR data for examining drivers of differential AD decline.

Existing approaches to identify AD populations from EHR have limitations. Using diagnosis codes alone yields poor accuracy due to lack of specificity for AD compounded by diagnostic delay.[17–20] Recent efforts using structured codified data and/or unstructured clinical information (*i.e.,* clinical narrative) largely relied on the EHR data of a single healthcare system and lacked external validation.[21–29] To validate the AD diagnosis phenotyping algorithms, prior studies used

labels such as clinical assessment, chart review, and surrogate indicators within the EHR.[21] Few AD-EHR cohorts hold the crucial linkage to AD research registries that contain gold-standard diagnosis labels for validating the algorithm performance. Finally, to our knowledge, no prior AD diagnosis phenotyping efforts validated performance across demographic groups despite the known differences in clinical presentations, which may exacerbate algorithm bias.

Here, we leveraged the EHR data of two independent healthcare systems, one of which is linked to an AD registry, applied a knowledge graph-guided unsupervised phenotyping algorithm to predict AD diagnosis across demographic groups, and validated algorithm performance using chart-reviewed and registry-derived gold-standard labels. Using the large real-world AD-EHR cohort, we then examined differences in two *readily ascertainable* clinical outcomes of AD decline (*i.e.,* nursing home admission, mortality) between demographic groups (*e.g.,* race and ethnicity, gender) and explored potential drivers of differential clinical decline in AD.

METHODS

Ethics Approval

The Institutional Review Board of the University of Pittsburgh (STUDY21020153) and Mass General Brigham (MGB, 2023P001450) approved the study protocol. The study using de-identified data is deemed exempt.

Harmonizing EHR data and creating AD-EHR data-marts

We obtained inpatient and outpatient EHR data from two healthcare systems, UPMC (Pittsburgh, PA; codified data: 2004-2022; narrative data: 2011-2022) and MGB (Boston, MA; codified and narrative data: 1994-2022). Codified data contain demographics, diagnoses (*e.g.*, International Classification of Disease [ICD] code), procedures (*e.g.*, Current Procedural Terminology [CPT] code), healthcare utilization metrics (*e.g.*, total number of ICD codes and encounter notes), medication prescriptions (*e.g.*, RxNorm codes), and laboratory tests (*e.g.*, Logical Observation Identifiers Names and Codes [LOINC] codes). We organized diagnoses by mapping ICD-9 and ICD-10 codes to PheCodes,[30,31] consolidated procedures using the Clinical Classifications Software for Services and Procedures (Agency for Healthcare Research and Quality), and mapped electronic prescriptions to RxNorm ingredient level. We deployed an established natural language processing pipeline (Narrative Information Linear Extraction [NILE]) to derive concept unique identifiers (CUIs) from clinical narratives (*e.g.,* physician office notes, discharge summaries) according to unified medical language system (UMLS).[32,33] Aiming high sensitivity to start, we built the *initial* **AD-EHR data-marts**, comprising UPMC (2011-2022) and MGB (1994-2022) patients with ≥1 ICD code for AD or related dementia (*e.g.,* ICD-9=290.x,

294.2x, 331.0; ICD-10=F03.9x, G30.x; **S-Table 1, S-Method-1**). We imputed missing race and ethnicity using age and gender.[34,35] We performed sensitivity analyses by excluding patients from the initial AD-EHR data-marts with missing race and ethnicity information (**S-Method-2**).

Identifying AD diagnosis

Using an unsupervised phenotyping algorithm, Knowledge driven Online Multimodal Automated Phenotyping (**KOMAP**),[36] we assigned AD diagnosis (*i.e.,* probable or possible AD vs *not* AD) for all patients in the initial AD-EHR data marts (**Figure 1A**). First, KOMAP leverages an online narrative and codified feature search engine (ONCE) powered by multi-source knowledge-graph representation learning to generate a list of informative codified and narrative features relevant to the target clinical concept, *i.e.,* AD (**S-Table 2**). Second, using the same list of ONCE-selected features, we trained KOMAP independently at UPMC and MGB to account for population heterogeneity. To validate KOMAP-predicted AD diagnosis, we used gold-standard chart-reviewed labels (UPMC, n=200; MGB, n=100, **S-Method-2**) and registry labels (University of Pittsburgh Alzheimer's Disease Research Registry [ADRC], n=1916 with UPMC EHR linkage). We reported standard performance measures (*e.g.,* area under the receiver operating characteristic [AUROC]) stratified by demographic groups.

Creating AD cohorts

We classified individuals from AD-EHR data-marts with KOMAP-predicted AD diagnosis at 90% specificity as having AD (**Figure 1B**). We held the specificity at 90% to minimize false positives while keeping reasonably large sample sizes. The first AD PheCode (*i.e.,* PheCode 290.11)

occurrence was the index date or *baseline*. For downstream analyses, we excluded patients with <24 months of EHR data or admitted to nursing homes prior to the baseline.

Covariates

We accounted for covariates, including demographics (*e.g.*, age at AD diagnosis, gender, race, ethnicity) and *baseline* clinical profiles (*e.g.,* healthcare utilization, pre-existing comorbidity burden) in the 24 months preceding the index date. Age at AD diagnosis was operationally defined as the age at the first AD PheCode. Gender included men and women. Race included American Indian or Alaska Native, Asian, Black, Native Hawaiian or Pacific Islander, White, and others. Ethnicity included Hispanic or Latino and non-Hispanic. We dichotomized race and ethnicity as NHW versus other minorities, which included individuals of Hispanic and non-European descent. Baseline healthcare utilization was the annualized total number of diagnosis codes and clinical encounters.[37] Baseline comorbidity burden was the Elixhauser comorbidity index (ECI).[38] Using the R *comorbidity* package, we mapped diagnosis codes to the 29 pre-existing health conditions comprising the ECI, and considered each comorbidity as an individual covariate.[39]

Clinical outcomes of AD decline

We assessed two clinically relevant indicators of AD decline readily ascertainable from EHR, time to nursing home admission and time to death. Nursing home admission is routinely documented in clinical practice,[4,5] and defined as having ≥1 code for admission to or clinical encounters in any type of residential institution (*e.g.,* assisted living facility, long-term care facility, end-of-life care facility, **S-Table 3**). Time to nursing home admission was from the index date to

the first occurrence of any nursing home code. Death status in EHR was linked to the social security death index. Time to death was from the index date to death. For patients without reaching either endpoint, we used all available EHR data.

Statistical analysis

We used two-tailed t-tests or two-proportion Z-tests to compare differences in demographics and baseline clinical profiles (**S-Table 4**) and prevalence of AD clinical outcomes among demographic groups (**S-Method-3**). We estimated the time-to-outcome using healthcare system-wise covariate-adjusted Cox proportional hazards (PH) models, stratified by demographic groups. Using patient-level data from both sites, we performed a fixed-effects meta-analysis using inverse variance weighting to estimate the time-to-outcome stratified by demographic groups, adjusting for pooled covariates. (Random-effects models would be unstable with two sites).

Data and code availability

Codes for KOMAP and project analysis are available on Github (**S-Method-4**).[40,41]

RESULTS

AD diagnosis phenotyping algorithm performance

Using chart-reviewed labels to evaluate AD diagnosis phenotyping algorithm performance, KOMAP achieved a higher AUROC at MGB (n=100, 45% AD, AUROC=0.923) than UPMC (n=200, 49% AD, AUROC=0.854, **Table 1, Figure 1C-F, S-Table 5, S-Table 6, S-Result-1**). While KOMAP achieved higher AUROC in women at both healthcare systems (UPMC: women=0.888, men=0.782; MGB: women=0.965, men=0.855), there was inconsistent pattern between two systems for race and ethnicity groups (UPMC: NHW=0.856, other=0.854; MGB: NHW=0.925, other=1.000). Using registry-derived labels to evaluate algorithm performance at UPMC (n=1916, 47% AD, AUROC=0.835, **S-Table 7**), KOMAP achieved higher AUROC in women than men and higher in racial and ethnic minorities than NHW individuals (women=0.846, men=0.824; NHW=0.821, other=0.922). In sensitivity analysis (without imputation of missing race and ethnicity), the algorithm achieved similar performances (**S-Table 8**).

AD cohort characteristics

After applying criteria (*i.e.,* ≥24 months of EHR data, no pre-baseline nursing home admission), the combined KOMAP-identified AD cohort comprised 34,181 patients (62% women, 90% NHW, mean[SD]=80.39[9.28] years at AD diagnosis, **Table 2**), including 16,411 UPMC patients (64% women, 92% NHW, 81.18[8.98] years at AD diagnosis) and 17,770 MGB patients (61% women, 89% NHW, 79.65[9.50] years at AD diagnosis). AD populations in the two healthcare systems differed in characteristics. Compared to MGB patients, UPMC patients had older age at AD

diagnosis, greater proportion of women and NHW individuals, greater baseline healthcare utilization and pre-existing comorbidity burden, greater proportion of nursing home admission though lower proportion of death during study follow-up, lower proportion of AD-related medication prescriptions (particularly before AD diagnosis), and shorter follow-up duration (**Table 2**).

In the combined AD cohort, demographic groups differed in patient profiles (**Table 2**). At AD diagnosis, women were older than men, while NHW individuals were older than racial and ethnic minorities (mean[SD] years: women=81.01[9.18], men=79.35[9.35], $p<.001$; NHW=80.55[9.23], other=78.86[9.61], $p<.001$). The pre-existing comorbidity burden was overall high (mean[SD] ECI score: 8.79[9.46], **S-Figure 2**), with men having higher burden than women and racial and ethnic minorities having greater burden than NHW individuals (mean[SD] ECI score: women=8.00[9.10], men=10.10[9.89], $p<.001$; NHW=8.68[9.40], other=9.88[10.02], $p<.001$), respectively. Baseline healthcare utilization was higher in men than women and higher in racial and ethnic minorities than NHW individuals (mean[SD] healthcare utilization counts: women=140.51[170.56], men=154.85[193.54], $p<.001$; NHW=144.94[178.48], other=154.84[190.23], $p<.001$). While most patients were prescribed AD-related medications, women and racial and ethnic minorities were prescribed at lower rates than men and NHW individuals, respectively (n[%] AD-related prescriptions: overall=27,219[80%]; women=16,825[79%], men=10,394[81%], $p<.001$; NHW=24,671[80%], other=2,548[78%], $p=.003$). Within follow-up, 10,910 (32%) patients were admitted to nursing homes and 16,928 (50%) died. A higher proportion of women and racial and ethnic minorities were admitted to nursing homes and died during follow-up than men and NHW individuals (n[%] admitted to nursing homes: women=6,967[33%], men=3,943[31%], $p<.001$; NHW=9,781[32%], other=1,129[34%], $p=.001$; n[%] death: women=10,931[51%], men=5,997[47%], $p<.001$;

NHW=15,110[49%], other=1,818[55%], *p*<.001). In sensitivity analysis (without imputation of missing race and ethnicity), patient profiles were comparable to the main AD cohort (**S-Table 9, S-Result-2**).

Differential risk of nursing home admission and death

Using fixed-effects meta-analysis of the healthcare system-wise covariate-adjusted Cox-PH models, we compared the risk of two pragmatic clinical outcomes among demographic groups. First, NHW individuals had a lower risk of nursing home admission than racial and ethnic minorities (HR [95% CI]=0.825 [0.776-0.877], *p*<.001, **Figure 2, S-Table 10, S-Result-3**), while there was no significant gender differences (HR [95% CI]=1.008 [0.967-1.050], *p*=0.762). Second, NHW individuals had a higher risk of death than racial and ethnic minorities (HR [95% CI]=1.381 [1.283-1.487], *p*<.0001, **S-Table 11**), while women had a lower risk than men (HR [95% CI]=0.873 [0.837-0.910], *p*<.0001). Key factors increasing nursing home admission risk (**Figure 3A**) included being a racial or ethnic minority, older age at AD diagnosis, high baseline healthcare utilization, and high pre-existing comorbidity burden. The main drivers of increased death risk (**Figure 3B**) included being NHW, male, older age at AD diagnosis, low healthcare utilization, and high pre-existing comorbidity burden.

DISCUSSION

We report using a novel knowledge graph-guided unsupervised phenotyping algorithm (KOMAP) to create a large real-world AD cohort from two independent healthcare systems as an AD clinical research resource. The algorithm achieved *overall* robust predictive performance in identifying people with AD. Demographic groups in this EHR-based AD cohort exhibited differential risks of clinical decline. Racial and ethnic minorities had a higher risk of nursing home admission, while NHW individuals and men had a higher risk of death.

The AD diagnosis phenotyping algorithm, while performing well overall, showed modest differences in healthcare system- and demographic group-specific performances, all of which still achieved reasonable AUROCs. To avoid co-training, we trained the same KOMAP algorithm independently at UPMC and MGB. KOMAP performed better at MGB compared to UPMC, potentially due to differences in clinical practices, documentation standards, population characteristics, and EHR data quality between the two healthcare systems. At both healthcare systems, KOMAP performed better among women than men, possibly attributable to women having more available EHR data for algorithm than men for several related reasons: (1) women have higher incidence and prevalence of AD;[1–3,42] (2) the real-world AD cohort had older age at AD diagnosis than traditional epidemiological studies;[43,44] (3) women survive longer than men, including after AD diagnosis.[10] Men with AD may also have greater overlap with other related dementias.[45] There was no consistent pattern of performance differences by race and ethnicity. KOMAP performed equally well among racial and ethnic groups at UPMC, while KOMAP in racial and ethnic minorities outperformed NWH individuals at MGB. The different patterns of healthcare utilization between UPMC and MGB populations may be contributory. KOMAP deployment for

AD diagnosis in future healthcare systems will require healthcare system-specific validation. Despite modest algorithm performance differences, the final combined AD cohort will be valuable for generating real-world evidence, particularly across demographic groups.

In this real-world AD cohort, the risk of clinical endpoints readily ascertainable from EHR differed across demographic groups. While gender groups had comparable risk of nursing home admission, men had a higher risk of death than women, potentially attributable to the greater pre-existing comorbidity burden in men, which was consistent with their higher baseline healthcare utilization. Further, men with AD may have a greater AD symptom burden (not measured here), leading to more rapid decline.[45] Racial and ethnic minorities had a higher risk of nursing home admission than NHW individuals, possibly also attributable to their greater pre-existing comorbidity burden and baseline healthcare utilization. Given their socioeconomic disadvantage,[4,46] racial and ethnic minorities may possibly receive lower quality nursing home care. Conversely, NHW patients had a higher risk of death than racial and ethnic minorities, mirroring the broader pattern in “mortality crossover” among older adults.[6–9,47] Racial and ethnic minorities more likely remain undiagnosed, leading to underestimation of AD-related mortality.[46] Our study cannot address whether biological factors drive differences in clinical outcomes across demographic groups.[48]

Our study suggests older age at AD diagnosis and greater pre-existing comorbidity burden (predominantly cardiometabolic diseases, psychosis) as drivers of higher risks of both nursing home admission and death.[10–16] Given that healthcare utilization and comorbidities generally increase with age, older age of diagnosis may reflect diagnostic delay (rather than later disease onset), possibly explaining its role in worse outcomes. Greater pre-existing comorbidity burden compounds AD management challenges and increases comprehensive and specialized care

needs, which are inaccessible to many patients. Interestingly, high baseline healthcare utilization (likely reflecting greater baseline functional impairment) increased nursing home admission risk, whereas low healthcare utilization (possibly reflecting inadequate symptom and comorbidity management) increased death risk. Taken together, clinical investigations using real-world AD cohorts would benefit from using multiple interrelated endpoints and careful interpretations.

The study has several strengths. First, the AD diagnosis phenotyping algorithm has the following innovations that enhance its generalizability. The algorithm leverages pre-trained and publicly available knowledge graphs of interconnected EHR concepts to rapidly select informative features, a strategy superior to reliance on clinician experts while effectively reducing feature dimensionality. To deploy, the unsupervised algorithm does not require any gold-standard labels (*e.g.,* confirmed AD diagnosis), which would require clinical expertise and substantial efforts. (To *evaluate* algorithm performance, we compared the algorithm-predicted AD diagnosis to a small number of gold-standard chart-reviewed and registry-derived labels.) Second, we validated the algorithm performance in two independent healthcare systems and across demographic groups. The overall robust performance in identifying people with AD supports our AD cohort as a clinical research resource. Third, we further demonstrated the long-term follow-up of the real-world AD cohort by examining time to two pragmatic but clinically relevant endpoints, *i.e.,* nursing home admission and death. Although >40% of people with AD eventually reside in nursing homes, nursing home admission has been underutilized even in EHR-based AD studies.[49–52] While other key clinical outcomes such as cognition and function fluctuate over time, nursing home admission and death are concrete milestones readily ascertainable from the EHR. Finally, we quantified the extent to which key demographic factors and pre-existing comorbidities drive AD decline. Specifically, we comprehensively assessed common comorbidities using ECI, a well-validated measure of comorbidity burden.

Our study also has limitations. First, we grouped multiple racial and ethnic groups (*e.g.,* American Indian, Asian, Black, Hispanic) as one minority group due to the modest sample size of each. While statistically necessary, this approach may obscure meaningful group differences and limit the generalizability to these demographic groups. Second, in the first study of this AD cohort, we did not assess cognitive or functional decline. Given the sparse EHR documentation of cognition (*e.g.,* Mini Mental Status Exam, Montreal Cognitive Assessment) and function (*e.g.,* Functional Activities Questionnaire) status,[53] we are developing novel methods to impute these longitudinal measures for future studies. Third, this AD cohort may include probable and possible AD and various clinical presentations (*e.g.,* amnestic, nonamnestic). Additional methods will be necessary to subtype the heterogeneous AD population. Through linkage with ADRC (at UPMC), we will compare the algorithm-predicted AD diagnosis with biomarker-informed AD diagnosis in the future.

CONCLUSION

We validated a novel algorithm to identify people with AD from EHR and created a large real-world AD cohort with long follow-up in two independent healthcare systems. This cohort study underscored the demographic differences facing the risk of nursing home admission and death and explored potential drivers. Future research using such real-world cohort will examine the contribution of additional factors (*e.g.,* social determinants of health as their coding are gradually introduced into routine clinical care) beyond demographics, healthcare utilization and comorbidities, deploy formal causal inference approaches, and incorporate imputed cognitive, functional and frailty status to complement nursing home admission and death as outcomes. This research can inform targeted interventions to reduce risk of AD decline.

ACKNOWLEDGEMENTS

This study was supported by the National Institutes of Health under award numbers R01 NS098023 and R01 NS124882 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

TABLES

Table 1. Performance of the AD diagnosis phenotyping algorithm evaluated using chart-reviewed labels at UPMC and MGB across demographic groups.

	UPMC (n = 200)					**MGB (n = 100)**				
	All	**NHW [b]**	**Other [C]**	**Men**	**Women**	**All**	**NHW [b]**	**Other [C]**	**Men**	**Women**
AUROC	0.854	0.858	0.854	0.782	0.888	0.923	0.925	1.000	0.855	0.965
AUPRC	0.825	0.843	0.732	0.709	0.866	0.904	0.903	1.000	0.868	0.926
Sensitivity [a]	0.595	0.602	0.630	0.447	0.726	0.750	0.857	1.000	0.667	0.808
PPV [a]	0.839	0.849	0.786	0.768	0.865	0.868	0.889	1.000	0.857	0.913
NPV [a]	0.716	0.807	0.709	0.687	0.789	0.804	0.875	1.000	0.778	0.833

a. We calculated the sensitivity, PPV, and NPV at a specificity of 90%. (See Methods for rationale of holding specificity at 90%.)
b. NHW, non-Hispanic White individuals
c. Other, racial and ethnic minorities

Other Abbreviations: *AUPRC*, area under the precision-recall curve; *AUROC,* area under the receiver operating characteristic curve; *NPV,* negative predictive value; *PPV*, positive predictive value

Table 2. Cohort characteristics.

	UPMC + MGB	**UPMC**	**MGB**		**UPMC + MGB by Race and Ethnicity**			**UPMC + MGB by Gender**		
	Overall	**Overall**	**Overall**	**P-value**	**NHW [g]**	**Other [h]**	**P-value**	**Men**	**Women**	**P-value**
Number of Patients, N (%)	34,181 (100%)	16411 (100%)	17770 (100%)		30,898 (100%)	3,283 (100%)		12,827 (100%)	21,354 (100%)	
Age at AD diagnosis, mean (SD)	80.39 (9.28)	81.18 (8.98)	79.65 (9.50)	**<.001**	80.55 (9.23)	78.86 (9.61)	**<.001**	79.35 (9.35)	81.01 (9.18)	**<.001**
Gender, N (%)				**<.001**			**<.001**			**<.001**
Women	21,354 (62%)	10,559 (64%)	10,795 (61%)		19,158 (62%)	2,196 (67%)		0 (0%)	21354 (100%)	
Men	12,827 (38%)	5852 (36%)	6,975 (39%)		11,740 (38%)	1,087 (33%)		12827 (100%)	0 (0%)	
Race and ethnicity [a], N (%)				**<.001**			**<.001**			**<.001**
Non-Hispanic White	30,898 (90%)	15,105 (92%)	15,793 (89%)		30,898 (100%)	0 (0%)		11,740 (92%)	19,158 (90%)	
American Indian	24 (<0.1%)	8 (<0.1%)	16 (<0.1%)		0 (0%)	24 (0.7%)		9 (<0.1%)	15 (<0.1%)	
Asian	361 (1.1%)	65 (0.4%)	296 (1.7%)		0 (0%)	361 (11%)		138 (1.1%)	223 (1.0%)	
Black	2,045 (6.0%)	1,183 (7.2%)	862 (4.9%)		0 (0%)	2,045 (62%)		640 (5.0%)	1,405 (6.6%)	
Hawaiian or Pacific Islander	6 (<0.1%)	3 (<0.1%)	3 (<0.1%)		0 (0%)	6 (0.2%)		4 (<0.1%)	2 (<0.1%)	
Hispanic or Latino	556 (1.6%)	47 (0.3%)	509 (2.9%)		0 (0%)	556 (17%)		185 (1.4%)	371 (1.7%)	
Other	291 (0.9%)	0 (0%)	291 (1.6%)		0 (0%)	291 (8.9%)		111 (0.9%)	180 (0.8%)	
Baseline healthcare utilization (number of events) [b], mean (SD)	145.89 (179.66)	205.27 (214.97)	91.05 (114.67)	**<.001**	144.94 (178.48)	154.84 (190.23)	**<.001**	154.85 (193.54)	140.51 (170.56)	**<.001**
Baseline Elixhauser comorbidity index [b], mean (SD)	8.79 (9.46)	9.04 (9.52)	8.56 (9.40)	**<.001**	8.68 (9.40)	9.88 (10.02)	**<.001**	10.10 (9.89)	8.00 (9.10)	**<.001**
Baseline comorbidities (top 2 prevalent), N (%)										
Hypertension, uncontrolled	22,418 (66%)	12,052 (73%)	10,366 (58%)	**<.001**	19,927 (64%)	2,491 (76%)	**<.001**	8,252 (64%)	14,166 (66%)	**<.001**
Cardiac arrhythmia	12,341 (36%)	6,326 (39%)	6,015 (34%)	**<.001**	11,147 (36%)	1,194 (36%)	0.7	5,424 (42%)	6,917 (32%)	**<.001**
Survival, N (%)				**<.001**			**<.001**			**<.001**

Alive	16,928 (50%)	9,180 (56%)	8,073 (45%)		15,788 (51%)	1,465 (45%)		6,830 (53%)	10,423 (49%)	
Death	17,253 (50%)	7,231 (44%)	9,697 (55%)		15,110 (49%)	1,818 (55%)		5,997 (47%)	10,931 (51%)	
Time to death, median months (1st and 3rd quantiles)	32 (14, 58)	31 (15, 53)	33 (13, 60)	**<.001**	32 (14, 57)	37 (16, 69)	**<.001**	29 (12, 52)	35 (16, 63)	**<.001**
Nursing Home Admission, N (%)				**<.001**			**0.001**			**<.001**
Admitted	10,910 (32%)	7,962 (49%)	2,948 (17%)		9,781 (32%)	1,129 (34%)		3,943 (31%)	6,967 (33%)	
Not admitted	23,271 (68%)	8,449 (51%)	14,822 (83%)		21,117 (68%)	2,154 (66%)		8,884 (69%)	14,387 (67%)	
Time to nursing home admission, median months (1st and 3rd quantiles)	12 (1, 33)	9 (0, 30)	21 (7, 46)	**<.001**	12 (1, 33)	17 (2, 43)	**<.001**	13 (1, 33)	12 (1, 34)	**<.001**
AD-related medication [c], N (%)										
Not prescribed	6,962 (20%)	3,709 (23%)	3,253 (18%)	**<.001** [d]	6,227 (20%)	735 (22%)	**0.003** [d]	2,433 (19%)	4,529 (21%)	**<.001** [d]
Prescribed	27,219 (80%)	12,702 (77%)	14,517 (82%)		24,671 (80%)	2,548 (78%)		10,394 (81%)	16,825 (79%)	
Before AD diagnosis	21,393 (79%)	8,885 (70%)	12,508 (86%)	**<.001** [e]	19,363 (78%)	2,030 (80%)	0.2 [e]	8,275 (80%)	13,118 (78%)	**0.001** [e]
On or after AD diagnosis	5,826 (21%)	3,817 (30%)	2,009 (14%)		5,308 (22%)	518 (20%)		2,119 (20%)	3,707 (22%)	
Follow-up duration [f], median months (1st and 3rd quantiles)	80 (45, 123)	78 (48, 109)	81 (45,129)	**<.001**	80 (46, 123)	80 (45, 131)	**0.01**	78 (45, 119)	81 (46, 126)	**<.001**

a. We imputed missing race and ethnicity using multiple imputation based on age and gender.
b. We calculated the baseline healthcare utilization and comorbidity burden in the 24 months prior to AD diagnosis.
c. The complete list of AD-related medications is available in S-Table 3.
d. We used a two-sample t-test to compare the difference between the number of patients prescribed and not prescribed AD-related medication.
e. We used a two-sample t-test to compare the difference between the number of patients prescribed AD-related medication before vs after AD diagnosis.
f. We estimated the median follow-up duration using the reverse Kaplan-Meier survival analysis.
g. NHW, non-Hispanic White individuals
h. Other, racial and ethnic minorities

FIGURES

Figure 1. Schematic overview. (A) Study design. (B) Inclusion criteria. (C-D) Phenotyping algorithm performance at UPMC. (E-F) Phenotyping algorithm performance at MGB.

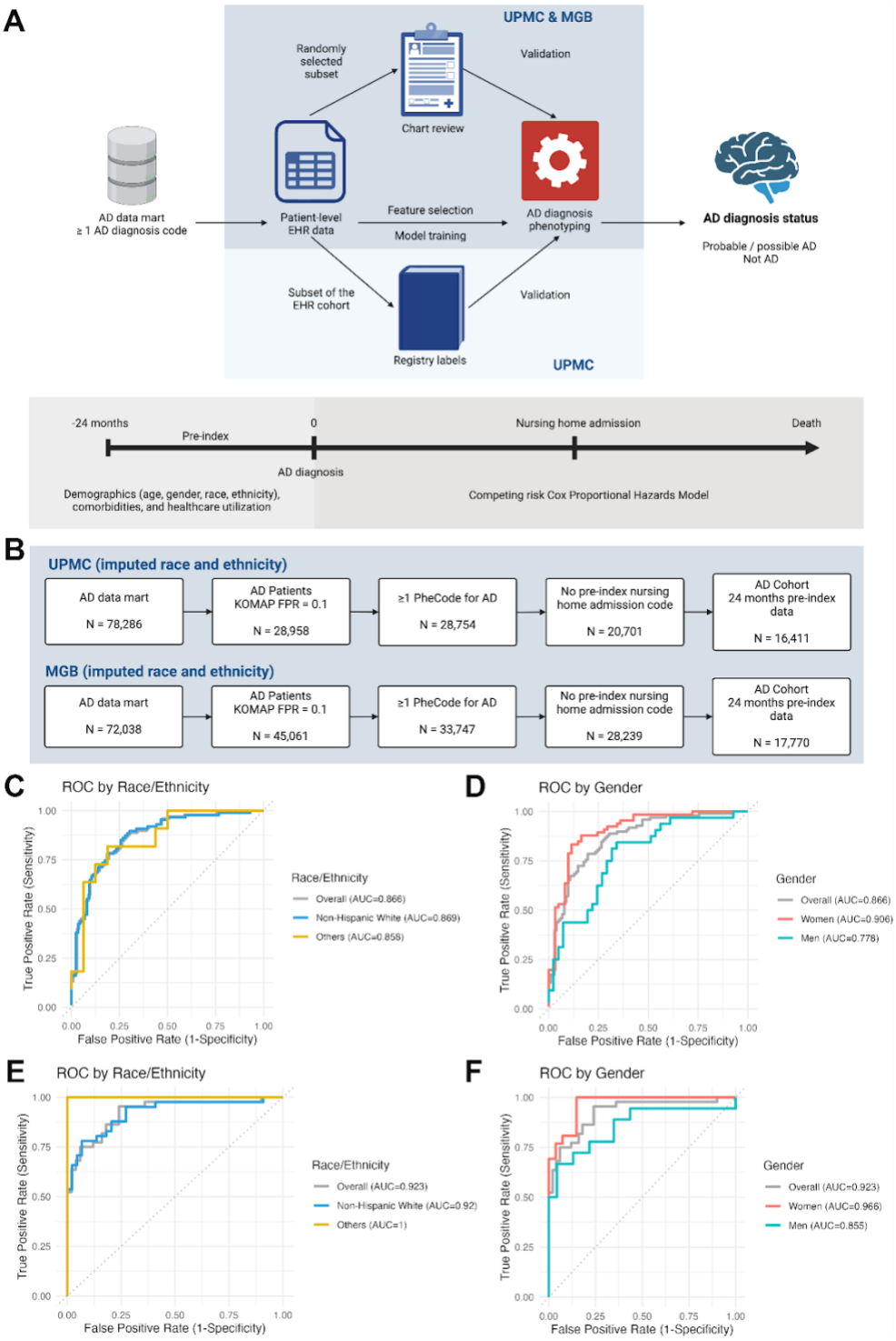


Figure 2. Fixed-effects meta-analysis of healthcare system-wise covariate-adjusted Cox proportional hazard models to estimate time to nursing home admission and death by demographic groups. The fixed-effects meta-analysis model was adjusted for pooled covariates across both healthcare systems from inverse-probability weighting. Covariates included demographics (*e.g.*, age at AD diagnosis, gender, race, ethnicity) and baseline clinical profiles (*e.g.,* healthcare utilization, pre-existing comorbidity burden) in the 24 months preceding the index date. (A-B) Nursing Home Admission (reference group: non-Hispanic White). (C-D) Death (reference group: women).

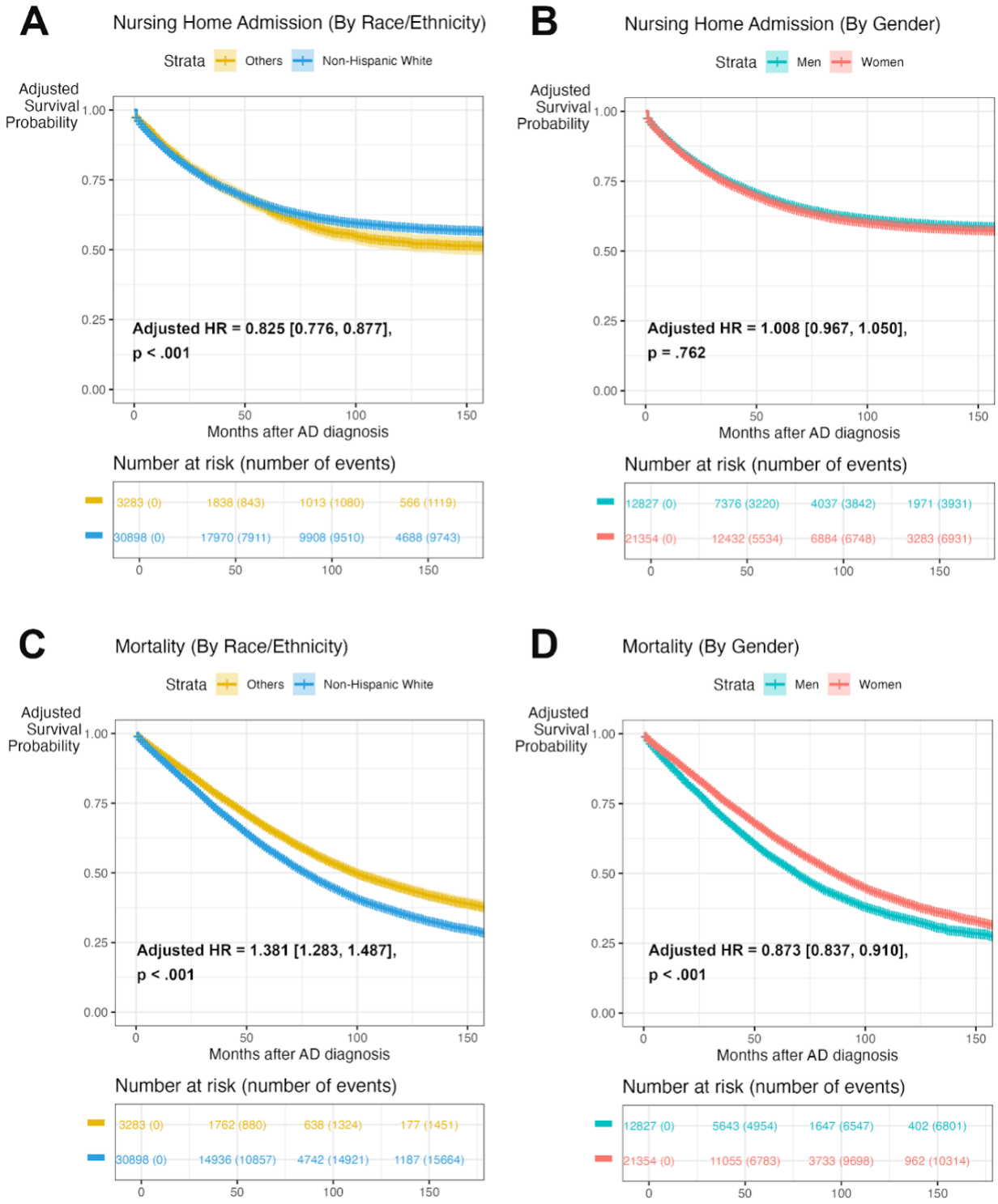
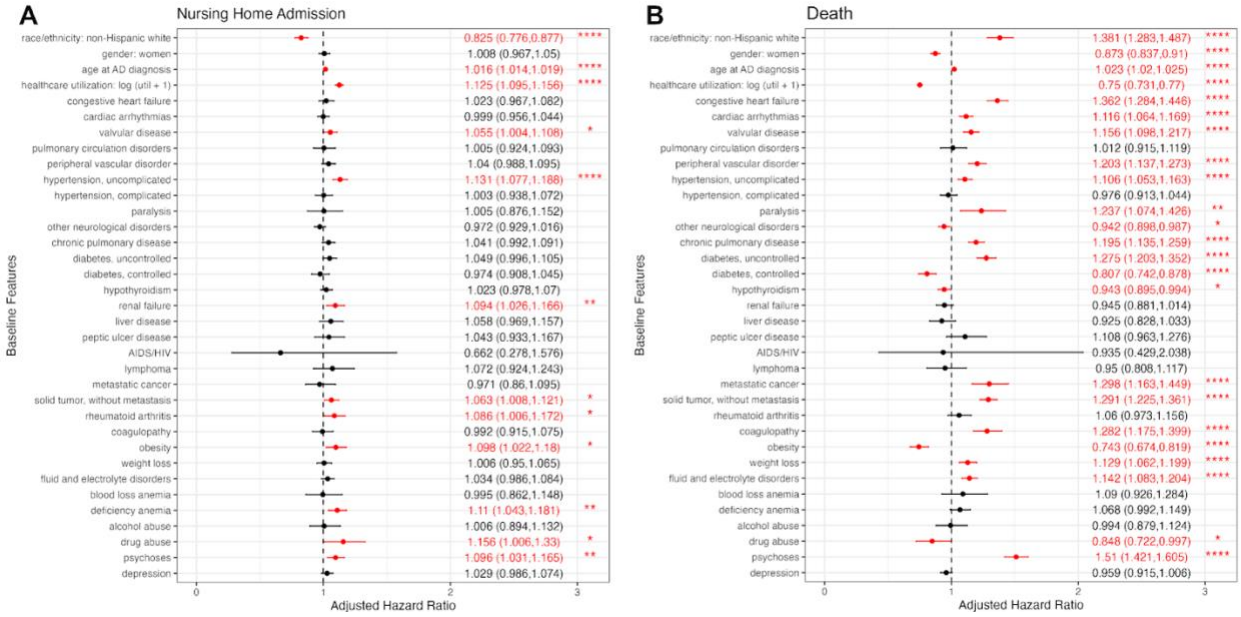


Figure 3. Adjusted hazard ratios of variables in the fixed-effects meta-analysis of healthcare system-wise covariate-adjusted Cox proportional hazard models of nursing home admission and death. The fixed-effects meta-analysis model was adjusted for pooled covariates across both healthcare systems from inverse-probability weighting. Covariates included demographics (*e.g.*, age at AD diagnosis, gender, race, ethnicity) and baseline clinical profiles (*e.g.,* healthcare utilization, pre-existing comorbidity burden) in the 24 months preceding the index date. (A) Nursing Home Admission. (B) Death.



REFERENCES

1. Scheltens P, Strooper BD, Kivipelto M, et al. Alzheimer's disease. *Lancet*. 2021;397(EMBO J 36 2017):1577-1590. doi:10.1016/s0140-6736(20)32205-4

2. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's Dement*. 2024;20(8):5143-5169. doi:10.1002/alz.13859

3. 2024 Alzheimer's disease facts and figures. Alzheimer's Dement. Published online 2024. doi:10.1002/alz.13809

4. Rivera-Hernandez M, Kumar A, Epstein-Lubow G, Thomas KS. Disparities in Nursing Home Use and Quality Among African American, Hispanic, and White Medicare Residents With Alzheimer's Disease and Related Dementias. *J Aging Heal*. 2019;31(7):1259-1277. doi:10.1177/0898264318767778

5. Temkin-Greener H, Yan D, Wang S, Cai S. Racial disparity in end-of-life hospitalizations among nursing home residents with dementia. *J Am Geriatr Soc*. 2021;69(7):1877-1886. doi:10.1111/jgs.17117

6. Mayeda ER, Glymour MM, Quesenberry CP, Johnson JK, Pérez-Stable EJ, Whitmer RA. Survival after dementia diagnosis in five racial/ethnic groups. Alzheimer's Dement. 2017;13(7):761-769. doi:10.1016/j.jalz.2016.12.008

7. Chen Y, Crimmins E, Ferido P, Zissimopoulos JM. Racial/Ethnic Disparities in Length of Life after Dementia Diagnosis: an 18-Year Follow-up Study of Medicare Beneficiaries. *Lancet Reg Heal - Am*. 2022;8:100179. doi:10.1016/j.lana.2021.100179

8. Khan AH, Ijaz E, Ubaid B, et al. Analysis of Alzheimer's Disease-Related Mortality Rates Among the Elderly Populations Across the United States: An Analysis of Demographic and

Regional Disparities from 1999 to 2020. *Curr Alzheimer Res*. 2024;21(6):384-394. doi:10.2174/0115672050338833240924113200

9. Zheng X, Wang S, Huang J, Li C, Shang H. Predictors for survival in patients with Alzheimer's disease: a large comprehensive meta-analysis. *Transl Psychiatry*. 2024;14(1):184. doi:10.1038/s41398-024-02897-w

10. Brück CC, Mooldijk SS, Kuiper LM, et al. Time to nursing home admission and death in people with dementia: systematic review and meta-analysis. *BMJ*. 2025;388:e080636. doi:10.1136/bmj-2024-080636

11. Yen TY, Beriwal N, Kaur P, et al. Medical Comorbidities and Association With Mortality Risk in Alzheimer's Disease: Population-Based Study of 132,405 Geriatric Inpatients. *Cureus J Medical Sci*. 2020;12(5):e8203. doi:10.7759/cureus.8203

12. Taudorf L, Nørgaard A, Brodaty H, Laursen TM, Waldemar G. Dementia increases mortality beyond effects of comorbid conditions: A national registry-based cohort study. *Eur J Neurol*. 2021;28(7):2174-2184. doi:10.1111/ene.14875

13. Rajamaki B, Hartikainen S, Tolppanen AM. The effect of comorbidities on survival in persons with Alzheimer's disease: a matched cohort study. *Bmc Geriatr*. 2021;21(1):173. doi:10.1186/s12877-021-02130-z

14. Armstrong MJ, Song S, Kurasz AM, Li Z. Predictors of Mortality in Individuals with Dementia in the National Alzheimer's Coordinating Center. *J Alzheimer's Dis*. 2022;86(4):1935-1946. doi:10.3233/jad-215587

15. Lanctôt KL, Hahn-Pedersen JH, Eichinger CS, et al. Burden of Illness in People with Alzheimer's Disease: A Systematic Review of Epidemiology, Comorbidities and Mortality. *J Prev Alzheimer's Dis*. 2024;11(1):97-107. doi:10.14283/jpad.2023.61

16. Sharma S, Liu J, Abramowitz AC, et al. Leveraging multi-site electronic health data for characterization of subtypes: a pilot study of dementia in the N3C Clinical Tenant. *JAMIA*

Open. 2024;7(3):ooae076. doi:10.1093/jamiaopen/ooae076

17. Phung TKT, Andersen BB, Høgh P, Kessing LV, Mortensen PB, Waldemar G. Validity of Dementia Diagnoses in the Danish Hospital Registers. *Dement Geriatr Cogn Disord*. 2007;24(3):220-228. doi:10.1159/000107084

18. Jr. DHT, Østbye T, Langa KM, Weir D, Plassman BL. The Accuracy of Medicare Claims as an Epidemiological Tool: The Case of Dementia Revisited. J Alzheimer's Dis. 2009;17(4):807-815. doi:10.3233/jad-2009-1099

19. Salem LC, Andersen BB, Nielsen TR, et al. Overdiagnosis of Dementia in Young Patients – A Nationwide Register-Based Study. Dement Geriatr Cogn Disord. 2013;34(5-6):292-299. doi:10.1159/000345485

20. Rizzuto D, Feldman AL, Karlsson IK, Aslan AKD, Gatz M, Pedersen NL. Detection of Dementia Cases in Two Swedish Health Registers: A Validation Study. J Alzheimer's Dis. 2018;61(4):1301-1310. doi:10.3233/jad-170572

21. Walling AM, Pevnick J, Bennett AV, Vydiswaran VGV, Ritchie CS. Dementia and electronic health record phenotypes: a scoping review of available phenotypes and opportunities for future research. *J Am Méd Inform Assoc*. 2023;30(7):1333-1348. doi:10.1093/jamia/ocad086

22. Oh IY, Schindler SE, Ghoshal N, Lai AM, Payne PRO, Gupta A. Extraction of clinical phenotypes for Alzheimer's disease dementia from clinical notes using natural language processing. JAMIA Open. 2023;6(1):ooad014. doi:10.1093/jamiaopen/ooad014

23. Estiri H, Azhir A, Blacker DL, Ritchie CS, Patel CJ, Murphy SN. Temporal characterization of Alzheimer's Disease with sequences of clinical records. *eBioMedicine*. 2023;92:104629. doi:10.1016/j.ebiom.2023.104629

24. Chen Z, Yang Y, Zhang D, et al. Predicting the Risk of Alzheimer's Disease and Related Dementia in Patients with Mild Cognitive Impairment Using a Semi-Competing Risk Approach. *Informatics*. 2023;10(2):46. doi:10.3390/informatics10020046

25. Li Q, Yang X, Xu J, et al. Early prediction of Alzheimer's disease and related dementias using real-world electronic health records. *Alzheimer's Dementia*. Published online 2023. doi:10.1002/alz.12967

26. Zhang Q, Coury R, Tang W. Prediction of conversion from mild cognitive impairment to Alzheimer's disease and simultaneous feature selection and grouping using Medicaid claim data. *Alzheimer's Res Ther*. 2024;16(1):54. doi:10.1186/s13195-024-01421-y

27. He X, Wei R, Huang Y, et al. Develop and validate a computable phenotype for the identification of Alzheimer's disease patients using electronic health record data. *Alzheimer's Dement: Diagn, Assess Dis Monit*. 2024;16(3):e12613. doi:10.1002/dad2.12613

28. Zhu W, Tang H, Zhang H, et al. Predicting Risk of Alzheimer's Diseases and Related Dementias with AI Foundation Model on Electronic Health Records. *medRxiv*. Published online 2024:2024.04.26.24306180. doi:10.1101/2024.04.26.24306180

29. Tang AS, Rankin KP, Cerono G, et al. Leveraging electronic health records and knowledge networks for Alzheimer's disease prediction and sex-specific biological insights. *Nat Aging*. 2024;4(3):379-395. doi:10.1038/s43587-024-00573-8

30. Liao KP, Sun J, Cai TA, et al. High-throughput multimodal automated phenotyping (MAP) with application to PheWAS. *J Am Med Inform Assn*. 2019;26(11):1255-1262. doi:10.1093/jamia/ocz066

31. Zhang Y, Cai T, Yu S, et al. High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). *Nat Protoc*. 2019;14(12):3426-3444. doi:10.1038/s41596-019-0227-6

32. Yu S, Cai T, Cai T. NILE: Fast Natural Language Processing for Electronic Health Records. *Arxiv*. Published online 2013. doi:10.48550/arxiv.1311.6063

33. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res*. 2004;32(suppl_1):D267-D270. doi:10.1093/nar/gkh061

34. Buuren S van, Groothuis-Oudshoorn K. mice : Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45(3). doi:10.18637/jss.v045.i03

35. Zhang G, Rose CE, Zhang Y, et al. Multiple Imputation of Missing Race and Ethnicity in CDC COVID-19 Case-Level Surveillance Data. *Int J Stat Méd Res*. 2022;11:1-11. doi:10.6000/1929-6029.2022.11.01

36. Xiong X, Sweet SM, Liu M, et al. Knowledge-Driven Online Multimodal Automated Phenotyping System. medRxiv. Published online 2023:2023.09.29.23296239. doi:10.1101/2023.09.29.23296239

37. Hou J, Kim N, Cai T, et al. Comparison of Dimethyl Fumarate vs Fingolimod and Rituximab vs Natalizumab for Treatment of Multiple Sclerosis. *Jama Netw Open*. 2021;4(11):e2134627. doi:10.1001/jamanetworkopen.2021.34627

38. Elixhauser* A, Steiner† C, Harris‡ DR, Coffey§ RM. Comorbidity Measures for Use with Administrative Data. *Méd Care*. 1998;36(1):8-27. doi:10.1097/00005650-199801000-00004

39. Gasparini A. comorbidity: An R package for computing comorbidity scores. *Journal of Open Source Software*. 2018;3(23):648. doi:10.21105/joss.00648

40. https://github.com/xialab2016/adphenotyping

41. https://github.com/xinxiong0238/KOMAP

42. Yan BW, Arias E, Geller AC, Miller DR, Kochanek KD, Koh HK. Widening Gender Gap in Life Expectancy in the US, 2010-2021. *JAMA Intern Med*. 2024;184(1):108-110. doi:10.1001/jamainternmed.2023.6041

43. Rentería MA, Mobley TM, Evangelista ND, et al. Representativeness of samples enrolled in Alzheimer's disease research centers. Alzheimer's Dement: Diagn, Assess Dis Monit. 2023;15(2):e12450. doi:10.1002/dad2.12450

44. Gianattasio KZ, Bennett EE, Wei J, et al. Generalizability of findings from a clinical sample to a community-based sample: A comparison of ADNI and ARIC. Alzheimer's Dement.

2021;17(8):1265-1276. doi:10.1002/alz.12293

45. Ferretti MT, Iulita MF, Cavedo E, et al. Sex differences in Alzheimer disease — the gateway to precision medicine. *Nat Rev Neurol*. 2018;14(8):457-469. doi:10.1038/s41582-018-0032-9

46. Hinton L, Tran D, Peak K, Meyer OL, Quiñones AR. Mapping racial and ethnic healthcare disparities for persons living with dementia: A scoping review. Alzheimer's Dement. 2024;20(4):3000-3020. doi:10.1002/alz.13612

47. Manton KG, Poss SS, Wing S. The Black/White Mortality Crossover: Investigation from the Perspective of the Components of Aging. *Gerontol*. 1979;19(3):291-300. doi:10.1093/geront/19.3.291

48. Research C on the U of R and E in B, Policy B on HS, Practice B on PH and PH, Services B on HC, Division H and M, Medicine NA of S Engineering, and. Rethinking Race and Ethnicity in Biomedical Research. Published online 2024. doi:10.17226/27913

49. Han E, Kharrazi H, Shi L. Identifying Predictors of Nursing Home Admission by Using Electronic Health Records and Administrative Data: Scoping Review. *JMIR Aging*. 2023;6:e42437. doi:10.2196/42437

50. Knapp M, Chua KC, Broadbent M, et al. Predictors of care home and hospital admissions and their costs for older people with Alzheimer's disease: findings from a large London case register. *BMJ Open*. 2016;6(11):e013591. doi:10.1136/bmjopen-2016-013591

51. Joling KJ, Janssen O, Francke AL, et al. Time from diagnosis to institutionalization and death in people with dementia. Alzheimer's Dement. 2020;16(4):662-671. doi:10.1002/alz.12063

52. Kosar CM, Mor V, Werner RM, Rahman M. Risk of Discharge to Lower-Quality Nursing Homes Among Hospitalized Older Adults With Alzheimer Disease and Related Dementias. *JAMA Netw Open*. 2023;6(2):e2255134. doi:10.1001/jamanetworkopen.2022.55134

53. Maserejian N, Krzywy H, Eaton S, Galvin JE. Cognitive measures lacking in EHR prior to dementia or Alzheimer's disease diagnosis. Alzheimer's Dement. 2021;17(7):1231-1243. doi:10.1002/alz.12280

54. Pinker E. Reporting accuracy of rare event classifiers. *npj Digit Med*. 2018;1(1):56. doi:10.1038/s41746-018-0062-0