Medicare Data for Dementia Research: Basics, Nuance, Access

Julie Bynum, MD MPH

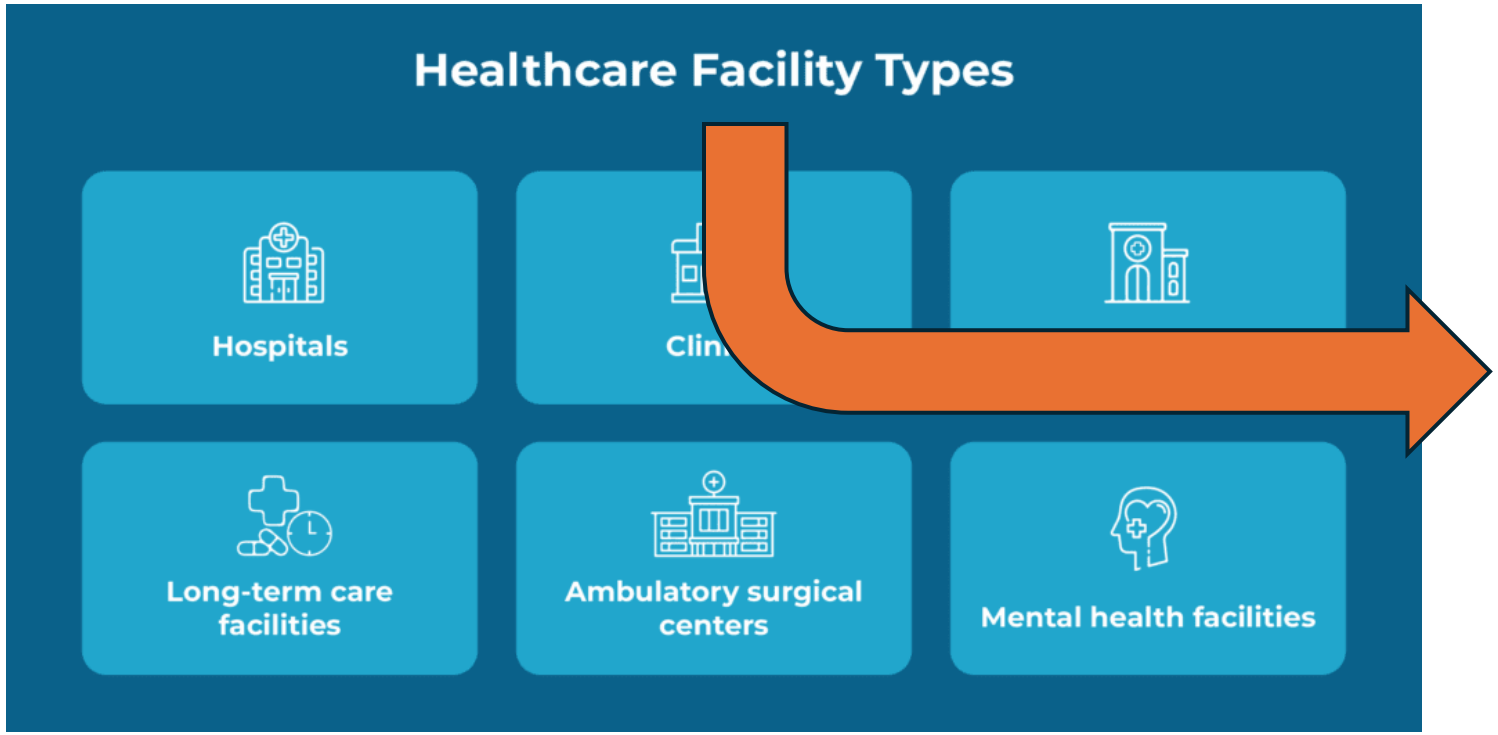
Professor of Internal Medicine
University of Michigan Medical School
November 12, 2025

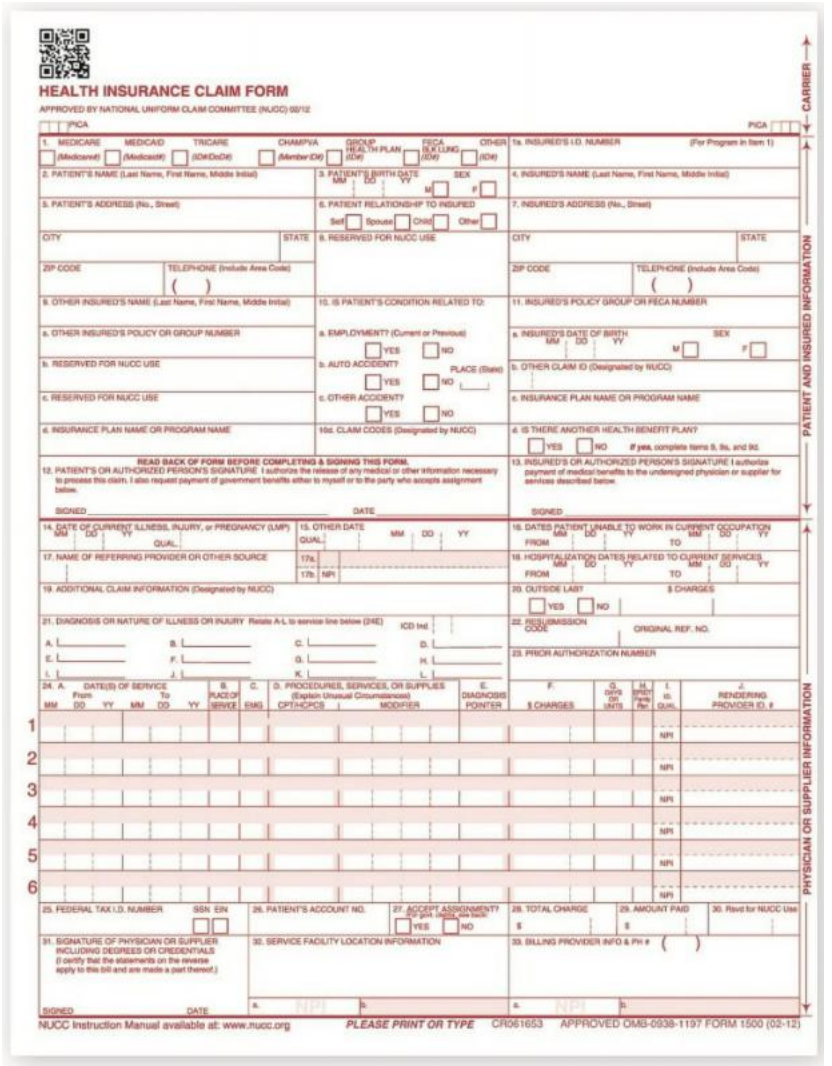
Healthcare-Generated Data

Data collected in the process of health care service delivery for payment or clinical record:

- ✓Medicare Fee-for-Service/Traditional Medicare (CMS)
- ✓Medicare Advantage (CMS)
- ✓Commercial Insurance (OPTUM, Sentinel/DRN, other payers)
- ✓Medicaid (CMS, state)
- ✓Minimum Dataset/OASIS (CMS)
- ✓Electronic Health Record

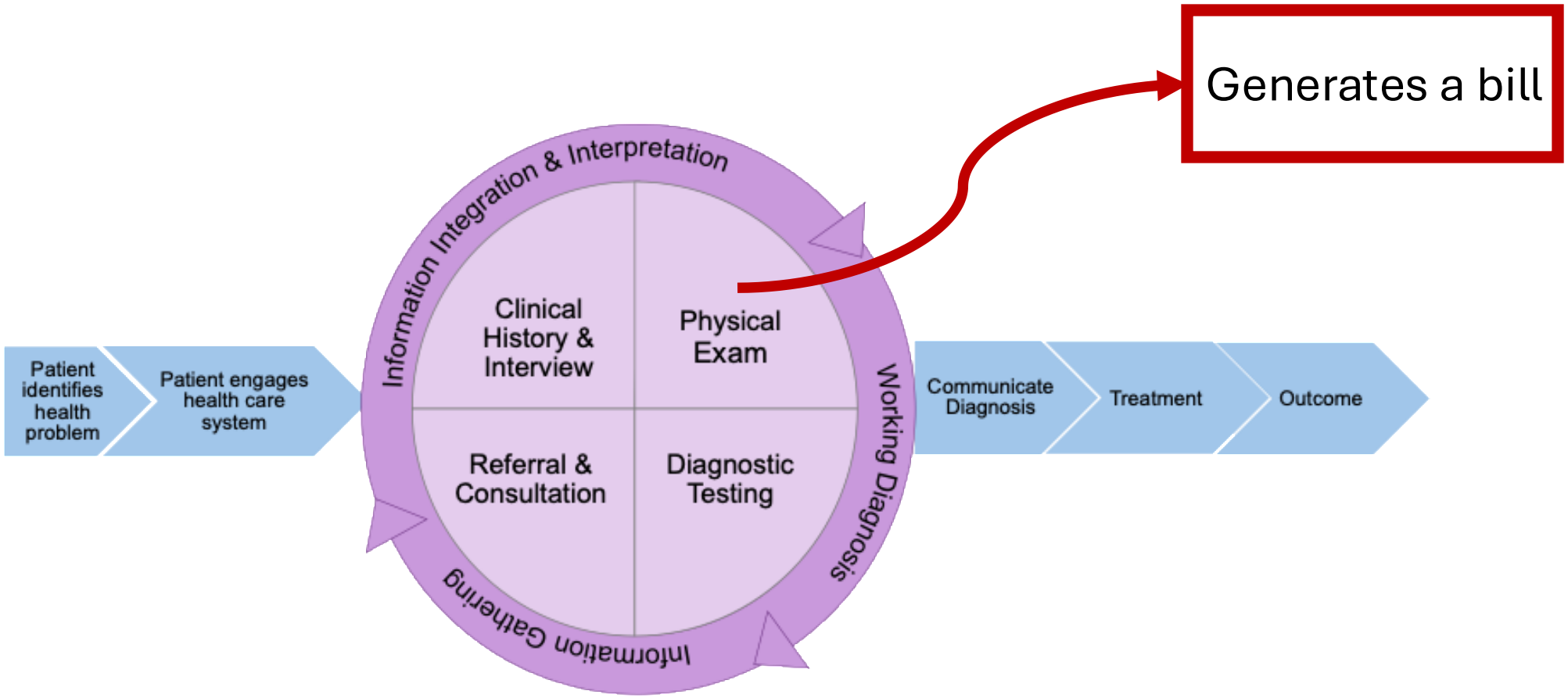
What is it Medicare Data?



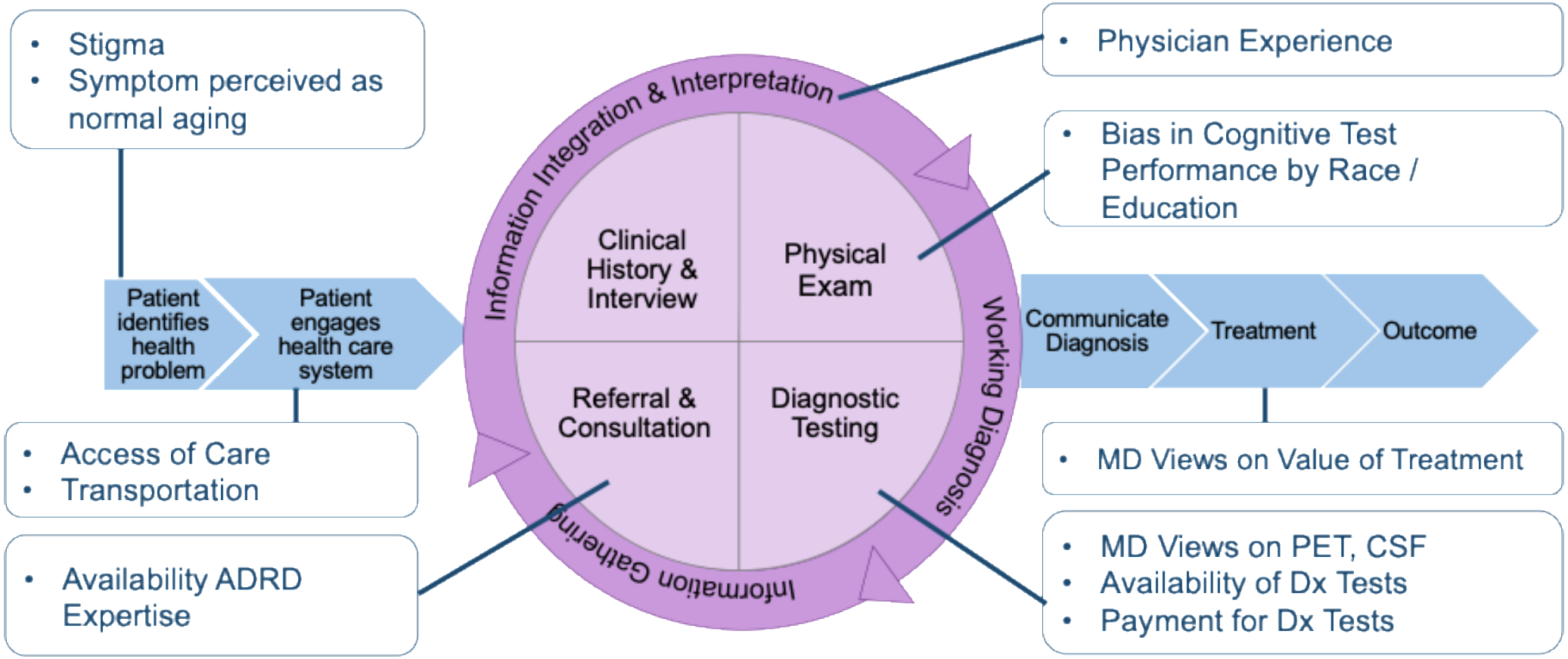


What is Medicare? Health insurance provided as an entitlement to people aged ≥65 who have paid 40 quarters of taxes & others based on disability or specific disease.

Reliance on Billed Service to Observe Cases



Reliance on Billed Service to Observe Cases



Conceptual Process of Diagnosis, NAM, 2014

Medicare Data Strengths & Challenges

Strengths

- Near complete age ≥ 65 population capture
- Uniform data elements across all healthcare sites
- Uniform data use agreement across all sites if CMS source
- Validated algorithms for dementia & other diseases

Weaknesses

Key Features of the Dementia Population at the National Level in Epidemiological Study vs. Traditional Medicare Claims

	OVERALL	AGE (Years)						SEX		RACE & ETHNICITY			
		65 -69	70-74	75-79	80-84	85-89	≥ 90	Female	Male	Black	Hispanic	White	Other
HCAP National Estimates (2016)* **Dementia % (95% CI)**	10% (9-11)	3% (1-4)	4% (2-6)	9% (6-11)	18% (14-22)	26% (20-31)	35% (28-43)	10% (9-11)	10% (8-11)	15% (10-19)	10% (7-13)	11% (10-13)	26% (13-39)
IMPACT Collaboratory 100% FFS Medicare (2019) **Dementia %****	6.2%	1.3%	2.5%	5.6%	10.8%	18.1%	27.0%	7.2%	5.1%	8.5%	7.2%	6.1%	4.5%

*Manly et al. In: *JAMA Neurology*. 2022: 1242-1249

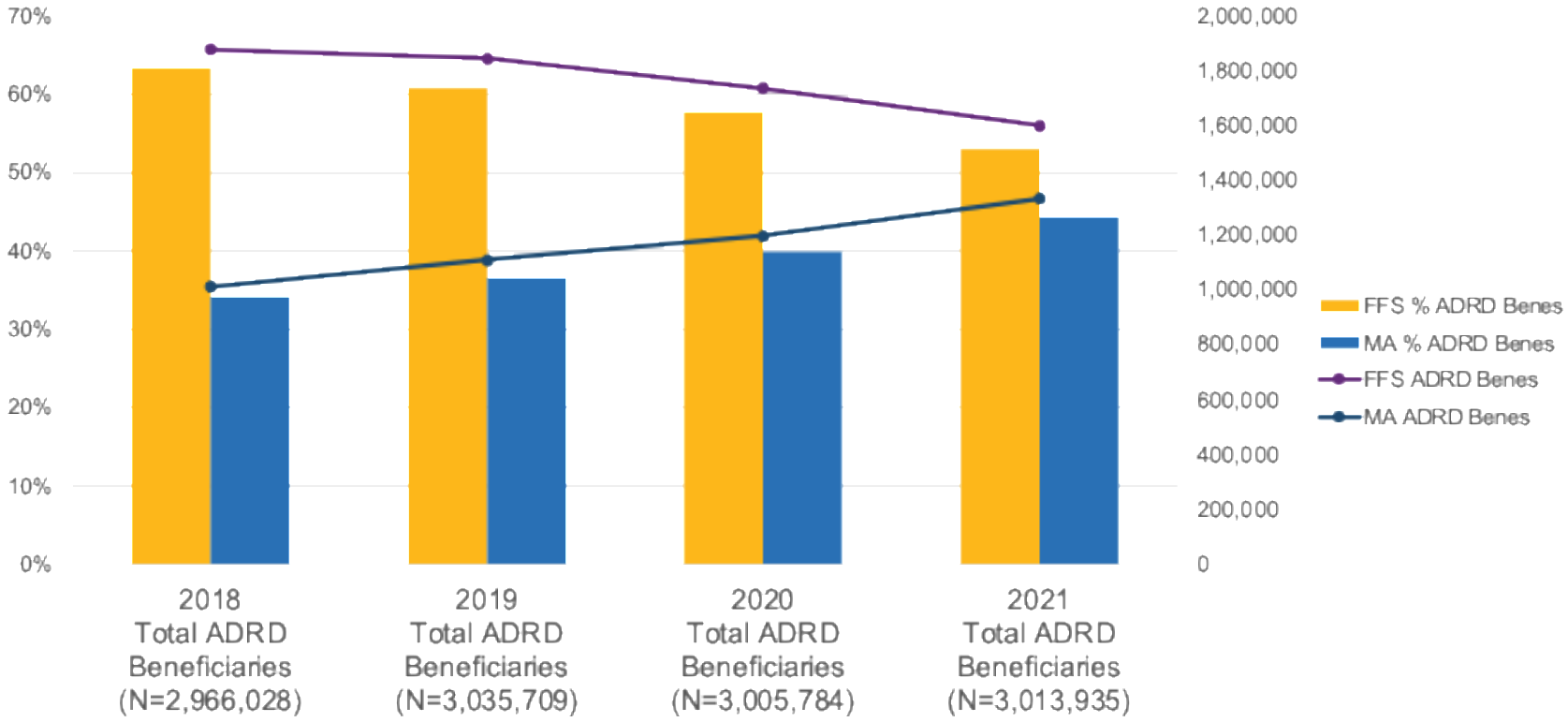
**IMPACT Collaboratory Data Analysis (2024)

Medicare Advantage – What's the difference?

In MA compared to TM

- Financial Incentives toward MORE diagnosis
- Enrollment influenced by disease
- Encounters not bill/claim

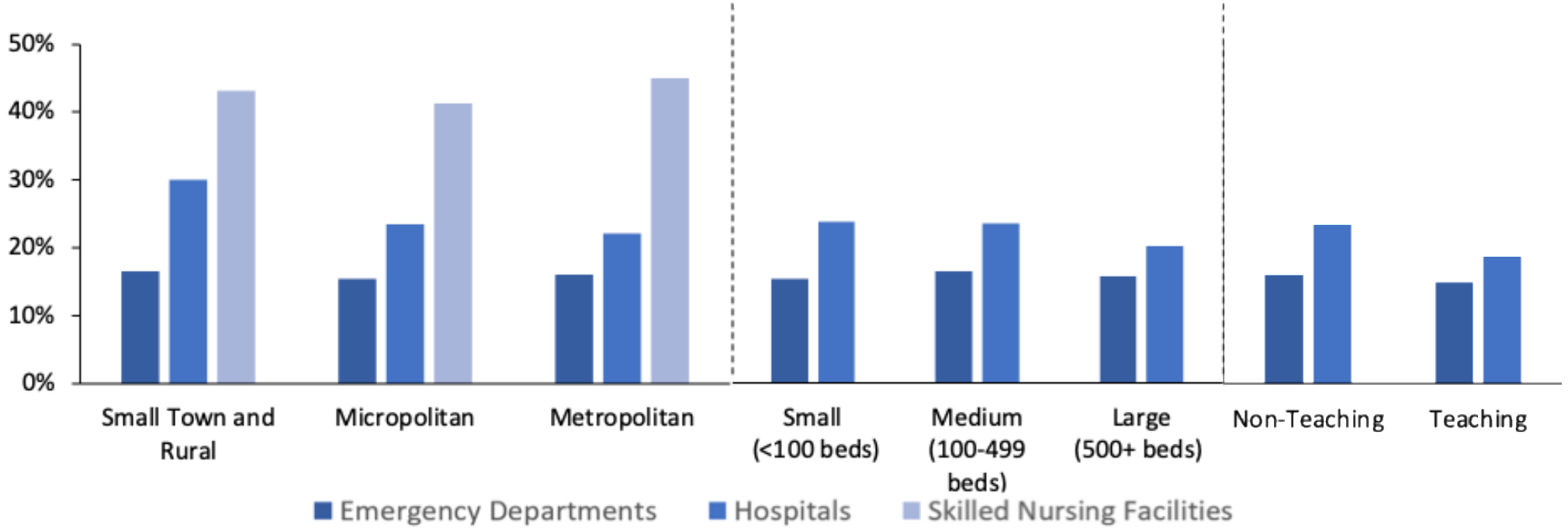
Number & Percent ADRD Beneficiaries TM vs. MA (2018-2021)



Setting of Care - ADRD Patients are Everywhere

ADRD Distribution across ED, Hospital, SNF:

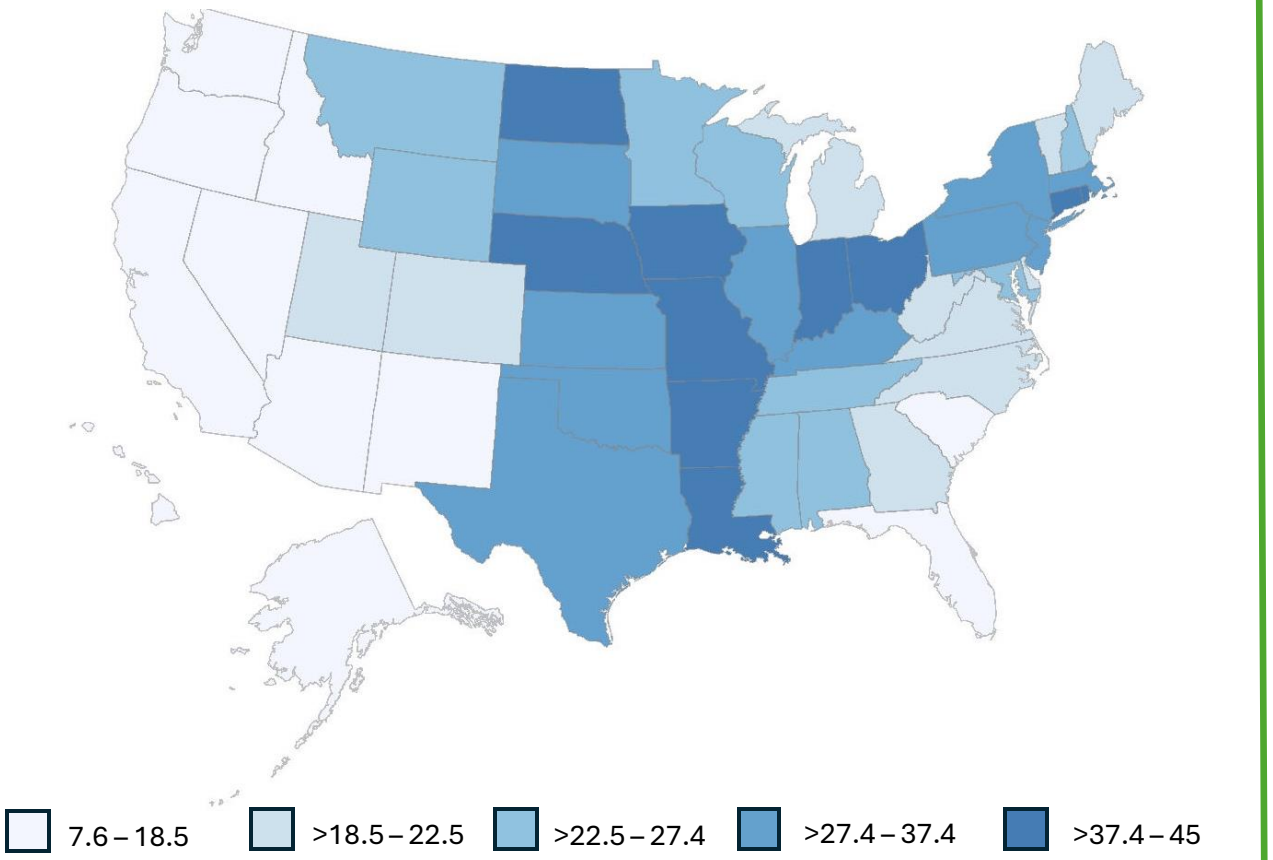
Mean Percent of Events from ADRD Patients Across Provider-Level Characteristics



Gettel, C., Martindale, J., & Bynum, J. (2024). *Health care use patterns among older adults with dementia*. CAPRA Data Brief. https://capra.med.umich.edu/data-briefs.html

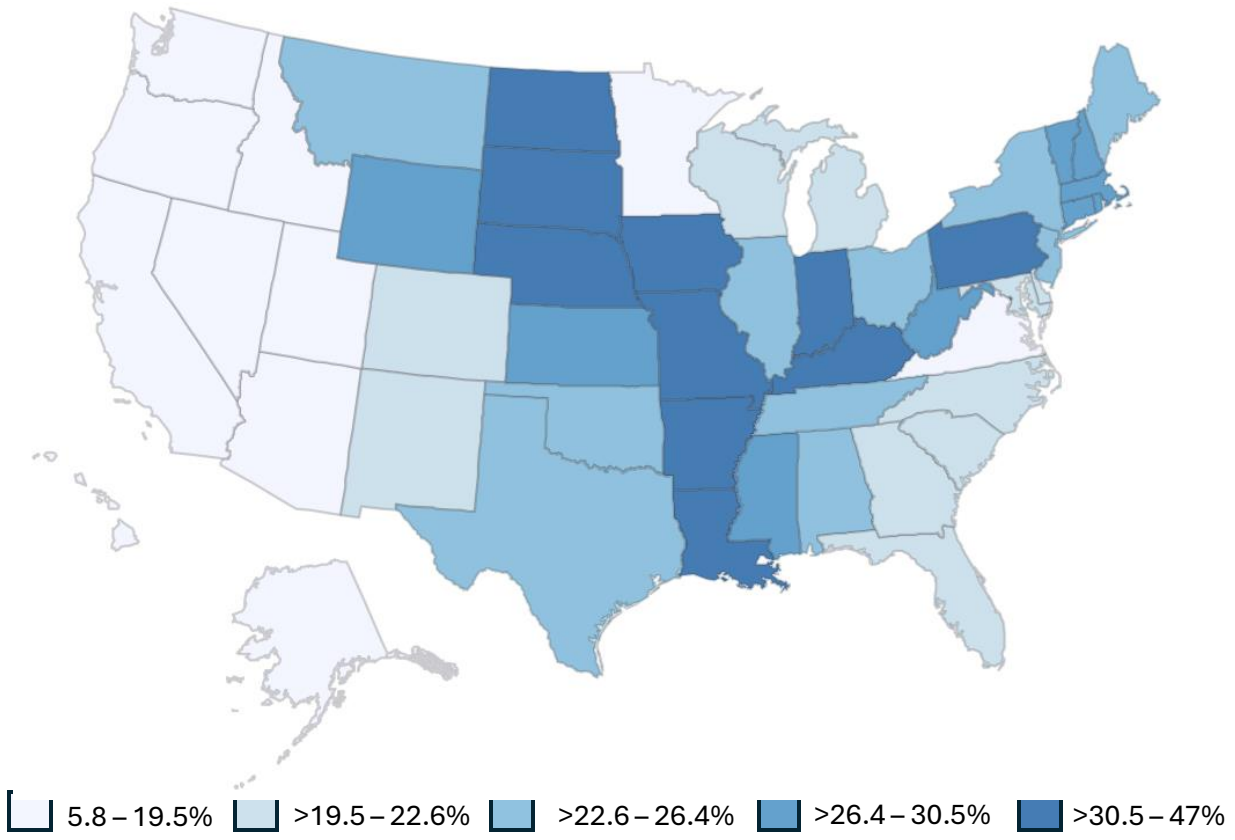
Nationally 24% Diagnosed ADRD Cases age 65+ in TM Reside in Nursing Homes (2019)

Certified NH Beds per 1000 Medicare FFS Enrollees (2019)



SOURCE: CMS, Office of Enterprise and Data Analytics, CMS Chronic Conditions Data Warehouse.

Percent Diagnosed ADRD in Long-stay Nursing Home (2019)

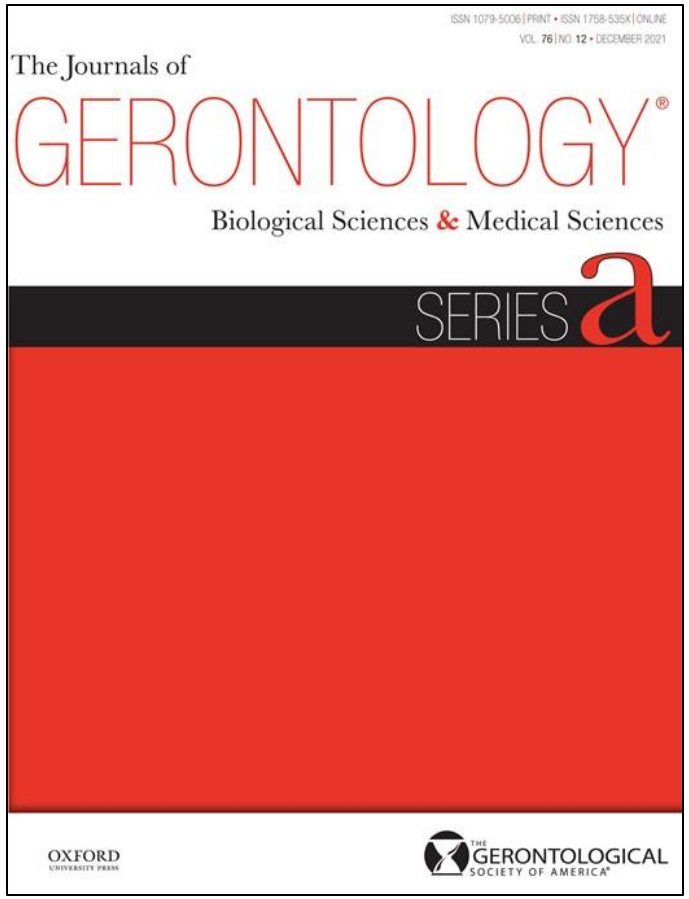


SOURCE: IMPACT Collaboratory Data Analysis (2024)

M | CAPRA

Diagnostic accuracy

Example of Participants Identified in Medicare Claims



Validation of Claims Algorithms to Identify Alzheimer's Disease and Related Dementias

Ellen P McCarthy, Ph.D., MPH, Chiang-Hua Chang, Ph.D., MS, Nicholas Tilton, Ph.D, Mohammed U Kabeto, MS, Kenneth M Langa, MD, Ph.D, Julie P W Bynum, MD, MPH

The Journals of Gerontology: Series A, glab373, https://doi-org.proxy.lib.umich.edu/10.1093/gerona/glab373

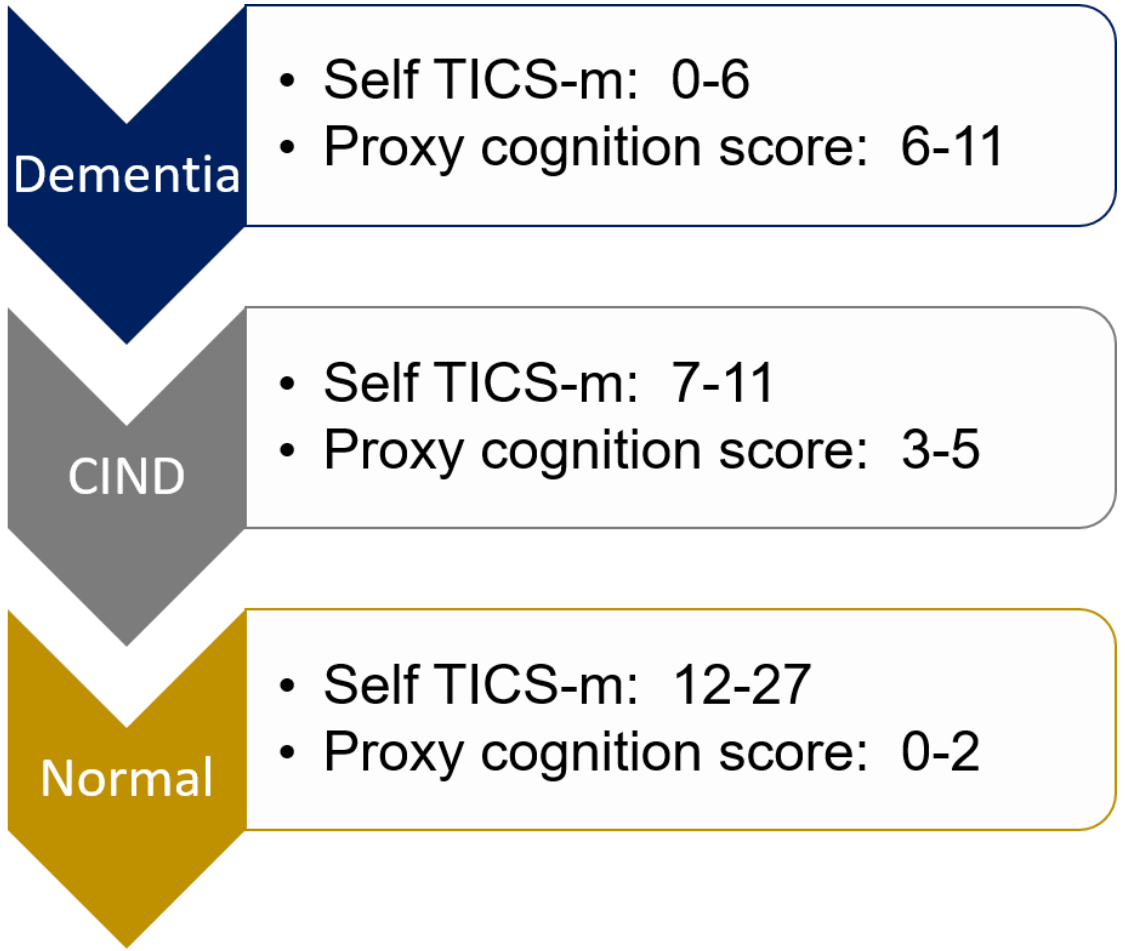
Identification Of Dementia In Recent Medicare Claims Data, Compared To Rigorous Clinical Assessments

Francine Grodstein, ScD ✉, Chiang-Hua Chang, PhD, Ana W Capuano, PhD, Melinda C Power, ScD, David X Marquez, PhD, Lisa L Barnes, PhD, David A Bennett, MD, Bryan D James, PhD, Julie P W Bynum, MD Author Notes

The Journals of Gerontology: Series A, glab377, https://doi-org.proxy.lib.umich.edu/10.1093/gerona/glab377

Reference Standard: HRS Cognitive Status

- Langa-Kebato-Weir method, cutpoints validated against ADAMS
  - Self: Modified TICS assessed immediate & delayed word recall, serial-7, and backward counting from 20
  - Proxy cognition score derived from proxy's assessment of participant's memory, difficulties with 5 IADLS, and interviewer rating of whether cognition was reason for proxy response

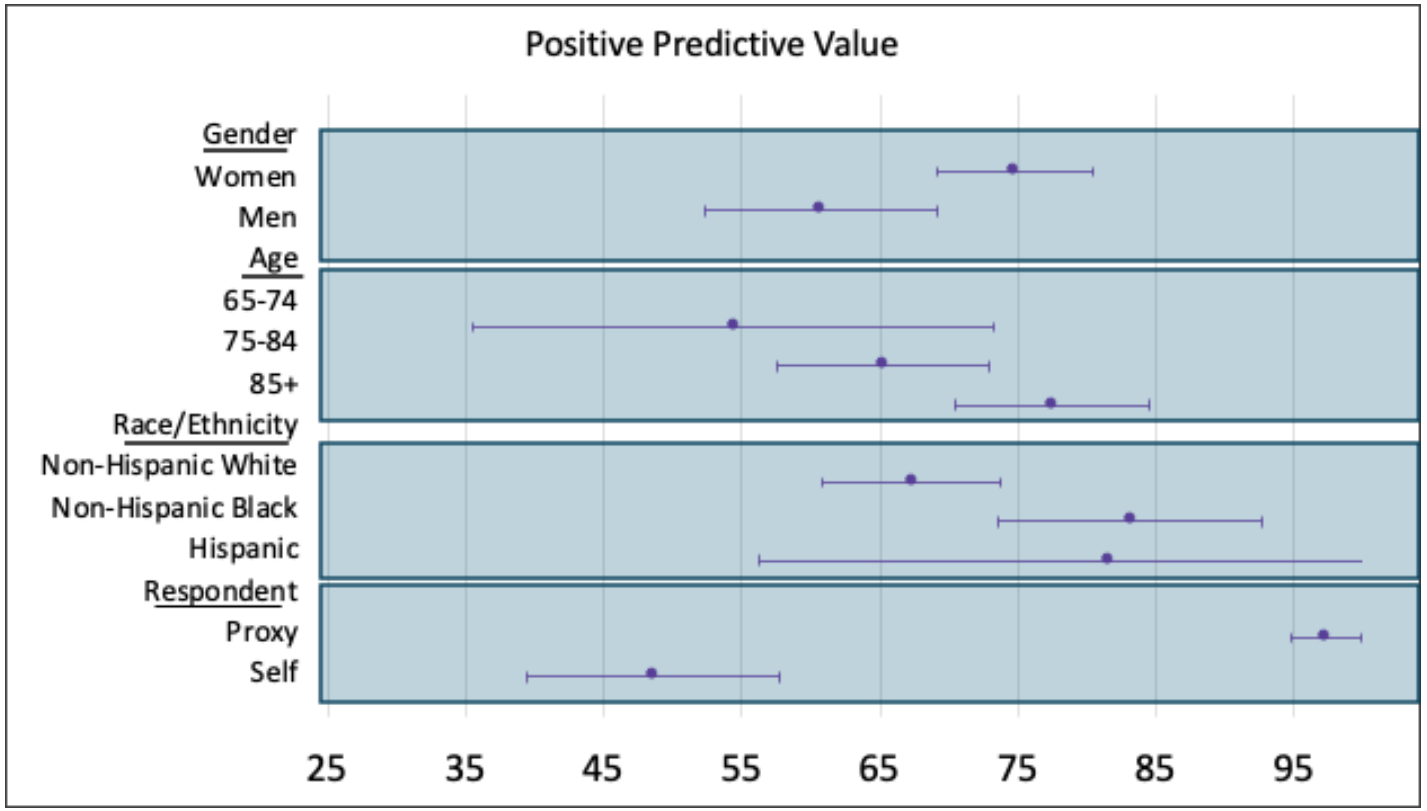


CIND: cognitive impairment not dementia
TICS-m: Modified Telephone Interview Cognitive Status

Standard 1-Year Algorithm Positive Predictive Value

Overall, claims had 70% PPV

(compared CCW 3yr algorithm 53% PPV)



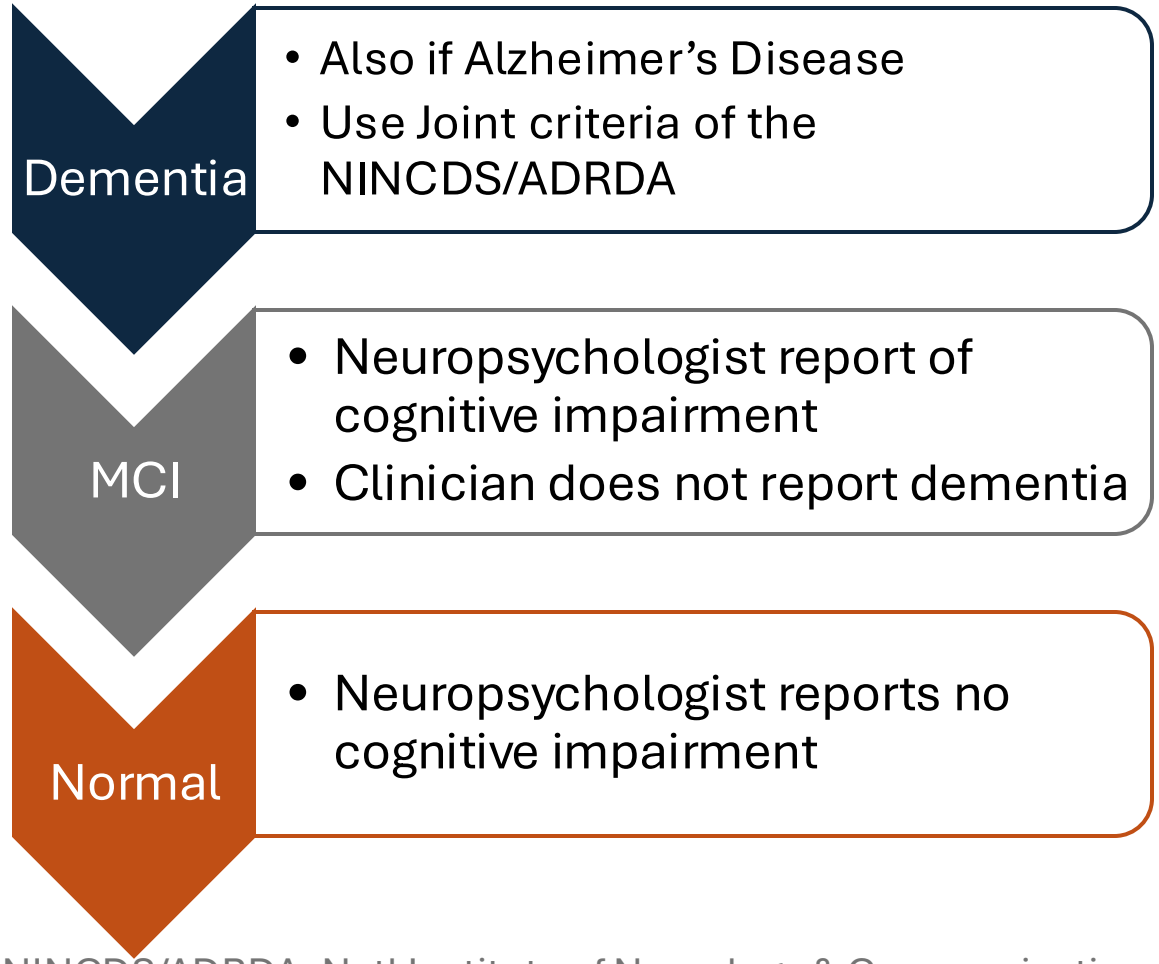
Positive Predictive Value is:

- Higher if person uses proxy in HRS
- Higher with older age
- Higher black or Hispanic (CI wide)

Gold Standard: Rush AD Cohorts Cognitive Status

- Annual, uniform, structured clinical evaluation harmonized across Cohorts
  - Neuropsychological and neurological evaluation
  - 18 cognitive tests
  - Severity rated across 5 domains
- Neuropsychologist blinded to sociodemographics renders clinical judgment on presence of dementia
- Experienced clinician reviews all data renders final clinical diagnosis

***5 Rush Cohort Studies**
Religious Orders Study
Rush Memory and Aging Project
Minority Aging Research Study
Rush African American Clinical Core
Rush Latino Core



NINCDS/ADRDA: Natl Institute of Neurology & Communication Disorders / Alzheimer's Disease & Related Dis Assoc.
MCI: Mild Cognitive Impairment

Who are the Dementia False Positives?

	DX in Cohort Eval		
DX in Claims	Yes	No	Total
Yes	92	66	158
No	52	908	960
Total	144	974	1118

Selected in, but <u>no dementia</u>

- Older
- More comorbidity
- More functional impairment
- Lower MMSE
- More MCI
- Frequent subjective complaints

False positives meet criteria for dementia in subsequent years:
@ 1 year 16%; @ 2 years 30%

	False Positives (N=66)	True Negatives (N=908)
Age	85 yrs	81 yrs
Male	24%	22%
White	78%	77%
Mean Education	16 yrs	17 yrs
MMSE	26.6	28.2
Subj Memory Concerns	51%	29%
Cohort DX MCI	72%	44%
ADL limitations, 3+	11%	5%
iADL limitations, 3+	34%	15%
Number Comorbidities	7	4
Hospitalized in year	59%	36%

$p < 0.001$

Who are the Dementia False Negatives?

	DX in Cohort Eval		
DX in Claims	Yes	No	Total
Yes	92	66	158
No	52	908	960
Total	144	974	1118

Missed cases, with dementia

- More likely Non-White
- Less functional impairment
- But no difference in subjective complaints or education

	False Negatives (N=52)	True Positives (N=92)
Age	89	90
Male	suppressed	19
White	75	90
Mean Educ	16	17
MMSE	20	15.4
Subj Memory Concerns	41	45
ADL limitations, 3+	suppressed	48
iADL limitations, 3+	52	82
Number Comorbidities	4	5
Hospitalized in year	39	56

$p < 0.001$

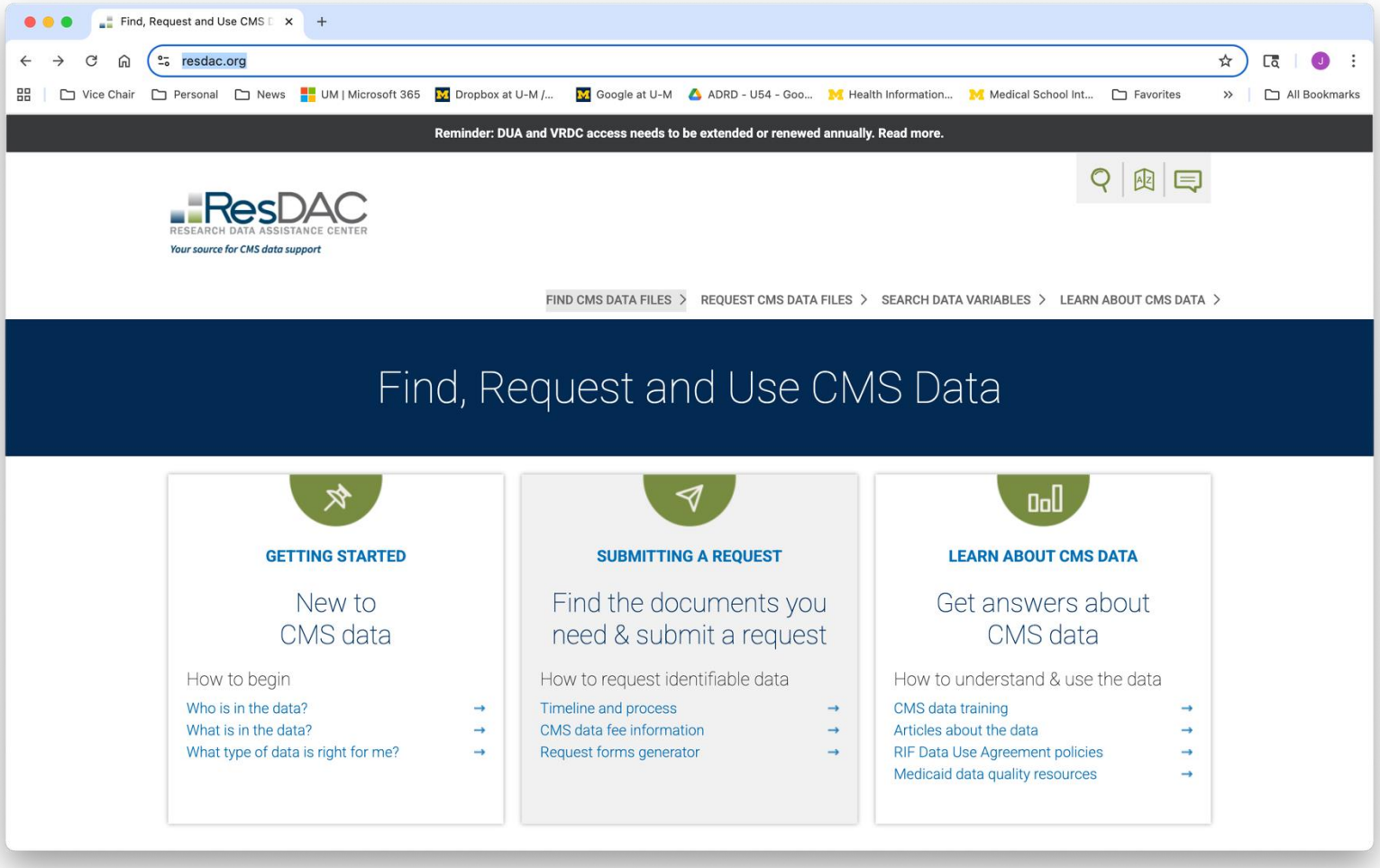
TM Claims-based ADRD Diagnostic Accuracy

Interpretation

- Sensitivity is the weakness of claims data
- Certain subgroups when flagged with ADRD are more likely to be accurately identified (older, uses a proxy, Black race, more severe disease)
- Subjective complaints common among false positives
- False positives are not normal cognitively or functionally
- False negatives more likely to be non-White and less functionally impaired

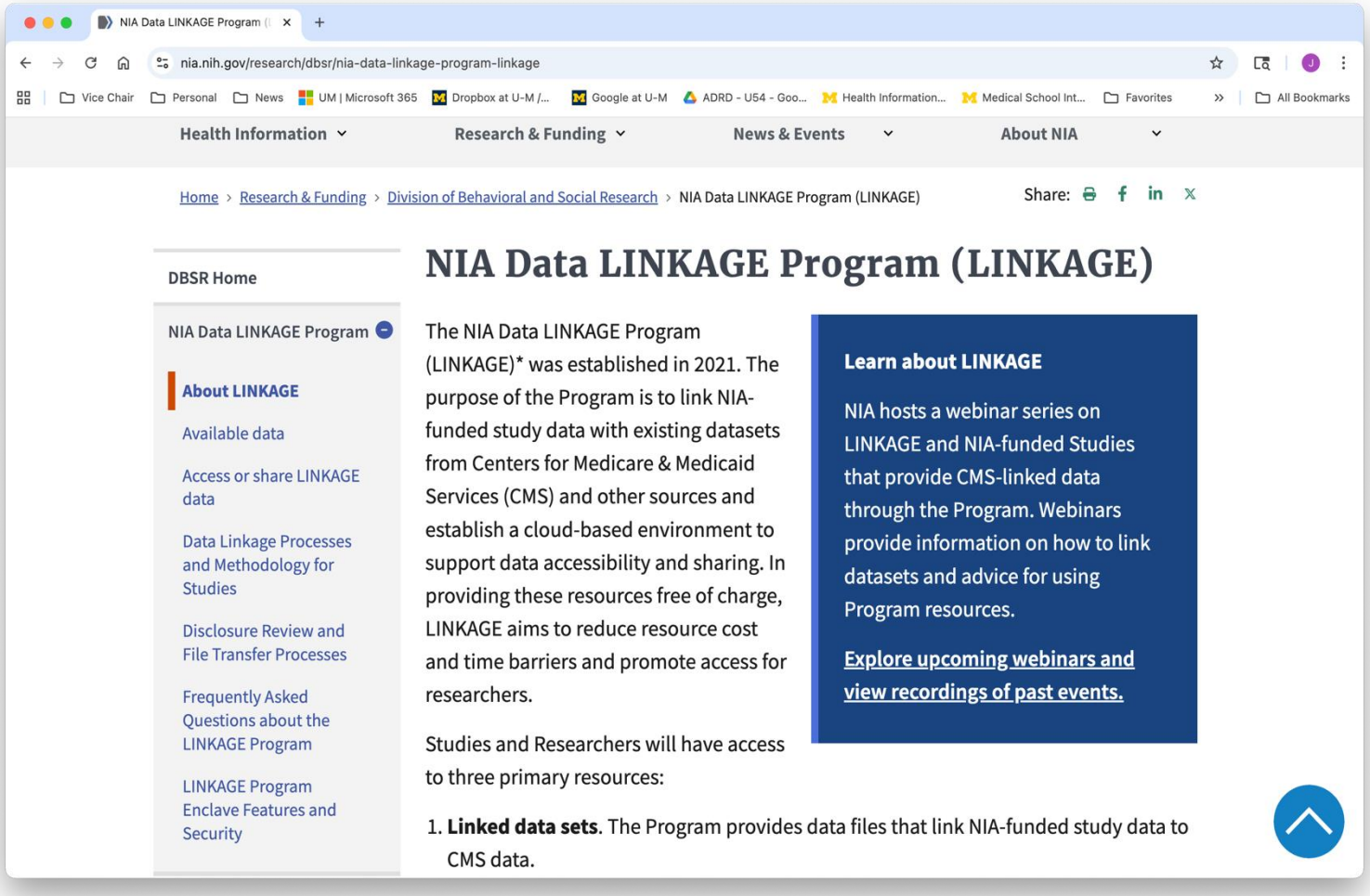
Note: MA Study coming soon

Access through CMS



- Stand alone CMS data
- Drawn from full population of Medicare enrolled
- Physical or enclave
- Application through ResDAC
- CMS Privacy Board

Access through NIA LINKAGE Program



- Based on NIA funded studies
- Study participants linked to their CMS data
- Enclave only
- Application to study and LINKAGE
- NIA Privacy Board

NIA Linkage Program
Studies Currently Linked to CMS Files

- Application of Economics & Social Psychology to Improve Opioid Prescribing Safety (AESOPS) Trial
- Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE)
- Boston Early Adversity & Mortality Study (BEAMS)
- Health and Retirement Study (HRS)
- LONG LIFE Family Study (LLFS)
- Long-Term Care Data (LTC) Cooperative
- Long-term Effects of a Community-based Volunteer Trial on Lifestyle Activity and Risk for Alzheimer's Disease (ExpCorps)
- Midlife in the United States (MIDUS)
- National Dementia Workforce Study (NDWS)
- National Health & Aging Trends Study (NHATS)
- National Long Term Care Survey (NLTCS)
- Optimizing Electronic Health Record Prompts with Behavioral Economics to Improve (NUDGE-EHR)
- Panel Study of Income Dynamics (PSID)
- Rush Alzheimer's Disease Center (RADC)
- Understanding America Study (UAS)
- Understanding Cohort Effects on Stroke, VCID, and Cognition After Major Epidemiologic Transitions. Previously REGARDS
- Wisconsin Longitudinal Study (WLS)

Summary

- Medicare data is available for nearly all people age 65+ in the U.S.
- These data have uniform access and data structure
- There is nuance to understanding who we see in these data relative to the true disease population
- Financial and other aspects of medical care and health policy influence what is observed and should not be forgotten when making inference
- Linkage to data with directly observed cognitive data provides opportunities for studies of dementia and cognitive impairment