

Subjective Cognitive Concerns and Risk for Cognitive Impairment in Longitudinal Electronic Health Records

Taiyang Chen, MS, Rebecca E. Amariglio, PhD, Ania Syrowatka, PhD, Li Zhou, MD, PhD, Gad A. Marshall, MD, Liqin Wang, PhD (lwang@bwh.harvard.edu)

Background

Subjective cognitive concerns (SCC), reported by patients, informants, or clinicians despite normal objective cognitive testing, may represent an early clinical marker of Alzheimer's disease and related dementias (ADRD).

In routine clinical care, SCC are typically documented only in free-text clinical notes rather than standardized diagnostic codes, limiting large-scale research and clinical applications.

Although SCC has been associated with future cognitive decline in research cohorts, its prognostic value in real-world electronic health records (EHRs) remains largely unknown.

Study Aims

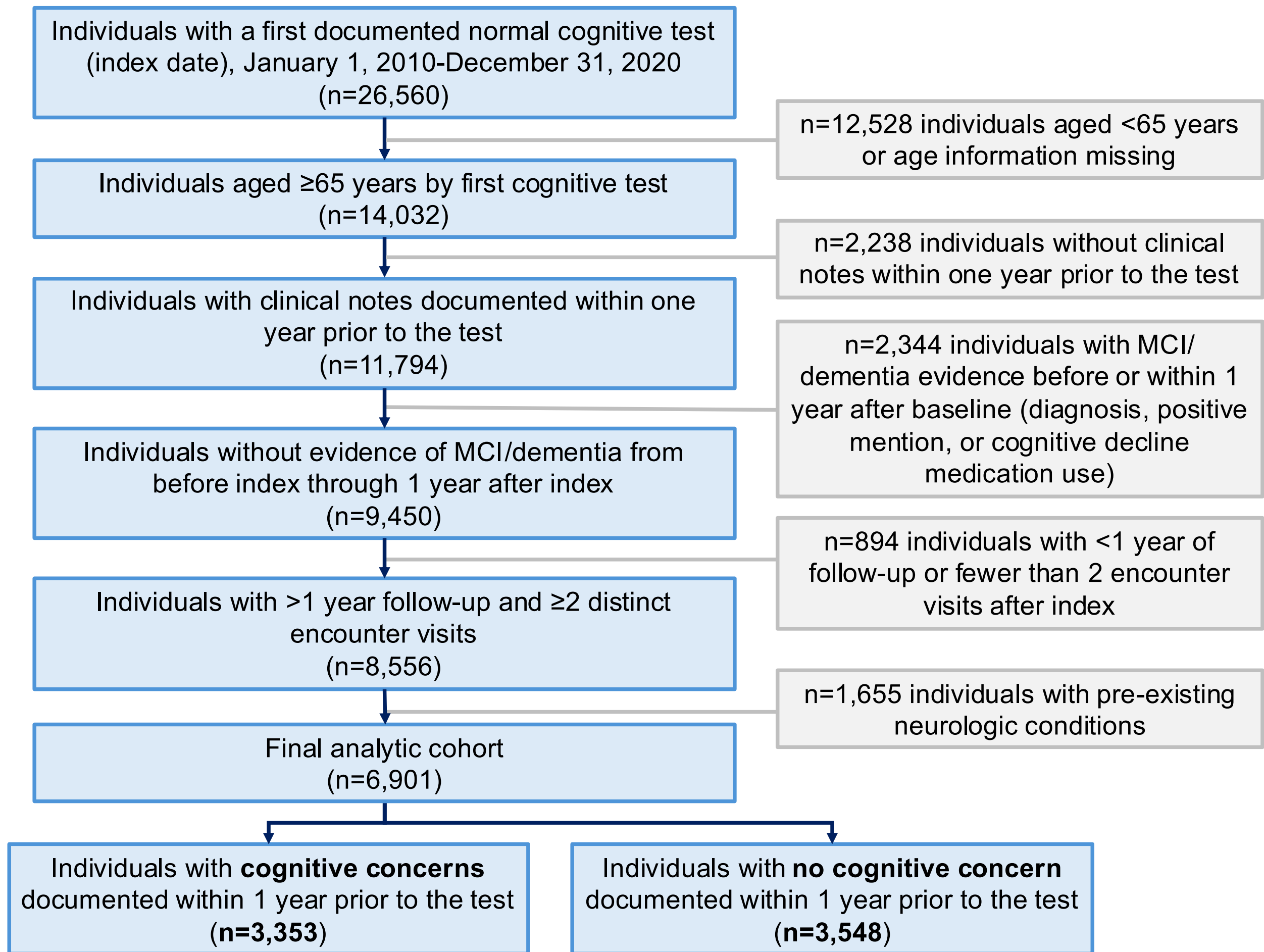
- Identify patient-, informant, or clinician-reported SCC from longitudinal EHR data using a large language model (LLM)-based pipeline.
- Compare baseline demographic and clinical characteristics between individuals with and without documented SCC.
- Estimate the association between documented SCC and incident mild cognitive impairment (MCI)/dementia using competing-risk regression models.

Results

Characteristic	No concern N = 3,548	Concern N = 3,353	p-value
Age, mean (SD) years	74 (6)	74 (7)	0.92
Sex, female	1,974 (56%)	1,952 (58%)	0.032
Conditions			
Hypertension	2,515 (71%)	2,284 (68%)	0.013
Coronary artery disease	999 (28%)	1,029 (31%)	0.021
Heart failure	228 (6.4%)	279 (8.3%)	0.003
Type 2 diabetes mellitus	550 (16%)	626 (19%)	<0.001
Hypothyroidism	591 (17%)	707 (21%)	<0.001
Vitamin B12 deficiency	108 (3.0%)	150 (4.5%)	0.002
Obesity	998 (28%)	658 (20%)	<0.001
Hyperlipidemia	2,001 (56%)	2,098 (63%)	<0.001
Depression	525 (15%)	1,143 (34%)	<0.001
Anxiety	540 (15%)	918 (27%)	<0.001
Excessive alcohol use	48 (1.4%)	128 (3.8%)	<0.001
Urinary tract infection	294 (8.3%)	450 (13%)	<0.001
Acute upper respiratory infection	196 (5.5%)	265 (7.9%)	<0.001
Hearing Impairment	607 (17%)	809 (24%)	<0.001
Visual Impairment	709 (20%)	875 (26%)	<0.001
Smoking	338 (9.5%)	402 (12%)	0.001

	No concern N = 3,548	Concern N = 3,353
No. of incident MCI/dementia events	746 (21)	1,899 (57)
No. of deaths	789 (22)	517 (15)
Follow-up time, years (mean, SD)	5.9 (3.2)	4.4 (3.9)
Incident rate per 100 person-years (95% CI)	3.5 (3.3, 3.8)	12.9 (12.3, 13.4)
Unadjusted Hazard ratio (HR) (95% CI)	1.00 (ref)	3.44 (3.16, 3.74)
Demographics-adjusted HR (95% CI)	1.00 (ref)	3.44 (3.16, 3.74)
Fully adjusted HR (95% CI)	1.00 (ref)	3.27 (2.99, 3.57)

Methods



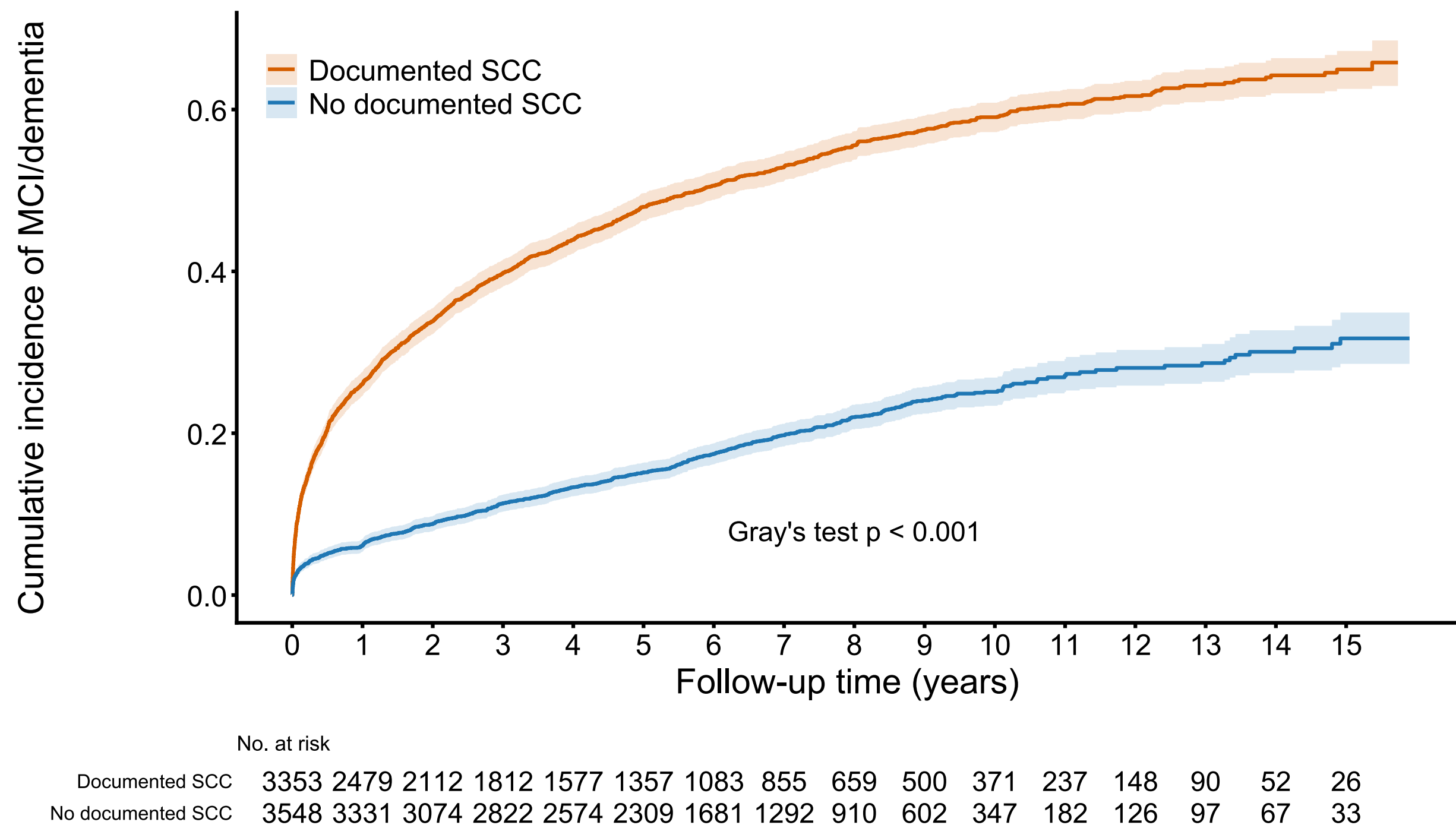
Study cohort

- N = 6,901** older adults with normal baseline cognitive test
 - Documented SCC: n = 3,353
 - No documented SCC: n = 3,548
 - Mean age: 74 years in both groups

Baseline differences

- Compared with the no-SCC group, individuals with SCC had a higher prevalence of:
 - Depression: 34% vs. 15% (p<0.001)
 - Anxiety: 27% vs. 15% (p<0.001)
 - Hypothyroidism, type 2 diabetes, hearing/visual impairment, alcohol use
- No SCC group had higher rates of obesity and hypertension

Main finding: Documented SCC were associated with a **3.27-fold higher risk** of incident MCI/dementia.



Study design: Retrospective cohort study using longitudinal EHR data from a large academic healthcare system (Mass General Brigham).

Population: Adults ≥65 years with a first documented normal cognitive test (MMSE, MoCA, Mini-Cog, SLUMS, or AD8) between January 1, 2010, and December 31, 2020. Cognitive test score and date were extracted using GPT-OSS 120B; extraction agreement on 200 manually annotated cases: κ = 0.911 and 0.903, respectively.

Exposure: Patient-, informant-, or clinician-documented SCC within 1 year prior to the index cognitive test, extracted from free-text clinical notes using GPT-OSS 120B (sensitivity 88.9%, PPV 84.2%, F1 86.5%).

Outcome: Incident MCI or dementia at least one year after the index date, identified using ICD diagnosis codes.

Statistical analysis:

- Fine-Gray competing-risk regression models (death as the competing event).
- Unadjusted, demographics-adjusted, and fully adjusted hazard ratios (HR) with 95% CIs
- Cumulative incidence curves by SCC status (Gray's test).
- Covariates: age, sex, race/ethnicity, marital status, education, prior-year encounter count, and comorbidities.

Conclusion

✓ Among older adults with normal baseline cognitive testing, documented subjective cognitive concerns were associated with more than a threefold higher risk of subsequent MCI or dementia, independent of demographic and clinical factors.

✓ SCC captured in routine EHR documentation, identified at scale using LLMs, may serve as an early, accessible, and clinically meaningful marker for identifying individuals at elevated risk of cognitive decline.

✓ These findings support further evaluation of scalable NLP-based SCC detection to improve risk stratification and facilitate earlier identification of older adults at increased risk of cognitive decline.

Acknowledgements

This work was supported by the Alzheimer's Association (AARF-22-924992) and the National Institute on Aging (R00AG075190).

