

Feasibility and Preliminary Validation of a Voice Agent-Administered Cognitive Screening Protocol in Healthy and Impaired Older Adults

Stephanie Ruth Young, MEd¹, PhD; Justin Mason, MBA² ;
¹Northwestern University Feinberg School of Medicine; ²CaringAI Inc.

P1-12

A Certified Dementia Practitioner accepted the AI screener’s triage recommendation in ~79% of cases ($\kappa = 0.72$), supporting the feasibility of telephone-based AI cognitive screening in primary care when paired with structured clinician oversight.

INTRODUCTION	METHODS	RESULTS	CONCLUSIONS	REFERENCES																														
Cognitive impairment is under-detected in primary care, and delays disproportionately affect older Black Americans with roughly twice the dementia rate of White Peers. Routine screening is limited by time, training, and staffing. AI voice agents enable scalable, structured screening by phone, familiar to older adults. We evaluated the scoring and recommendations of a telephone-based AI screening agent, CaringAI, via expert clinician audit, comparing in-clinic vs. scheduled outbound calls.	Sample. Older adult patients (n=68; ≥60 years old) at a family practice in an accountable care organization; predominantly Black/AA, Medicare/Medicaid-insured Screening. CaringAI Listen administered five instruments by phone - 6CIT, a supplementary cognitive set, Lawton IADL, GAD-2, and GDS-4. In-clinic calls at intake vs. scheduled outbound calls ~14 days before the visit. Review. A Certified Dementia Practitioner confirmed or overrode each AI score and triage recommendation. Analysis. ICC(2,1) for scores; Cohen’s κ for triage; bootstrap 95% CIs	AI-reviewer score agreement was good to excellent (ICC .84 to 1.00). Reviewer accepted 79.4% of AI triage recommendations AI-reviewer score concordance <table><tr><th>Instrument</th><th>ICC (95% CI)</th><th>Diff*</th></tr><tr><td>6CIT</td><td>0.84 (0.76–0.91)</td><td>–1.06</td></tr><tr><td>Supp. cognitive</td><td>0.85 (0.75–0.92)</td><td>–0.22</td></tr><tr><td>IADL</td><td>0.93 (0.72–1.00)</td><td>+0.10</td></tr><tr><td>GAD-2</td><td>0.93 (0.82–1.00)</td><td>0.00</td></tr><tr><td>GDS-4</td><td>1.00 (1.00–1.00)</td><td>0.00</td></tr></table> *Diff = reviewer – AI score (Bland-Altman mean). ICC (Koo & Li): >0.90 excellent; 0.75–0.90 good. Triage recommendations accepted without modification, by deployment mode <table><tr><th>Deployment Mode</th><th>Triage accepted (%)</th><th>κ</th></tr><tr><td>Overall (n=68)</td><td>79.4%</td><td>0.72</td></tr><tr><td>Sched. outbound (n=30)</td><td>86.7%</td><td>0.83</td></tr><tr><td>In-clinic (n=38)</td><td>73.7%</td><td>0.56</td></tr></table> Between-mode difference not significant: Δκ = 0.28 (95% CI -0.00 to 0.53)	Instrument	ICC (95% CI)	Diff*	6CIT	0.84 (0.76–0.91)	–1.06	Supp. cognitive	0.85 (0.75–0.92)	–0.22	IADL	0.93 (0.72–1.00)	+0.10	GAD-2	0.93 (0.82–1.00)	0.00	GDS-4	1.00 (1.00–1.00)	0.00	Deployment Mode	Triage accepted (%)	κ	Overall (n=68)	79.4%	0.72	Sched. outbound (n=30)	86.7%	0.83	In-clinic (n=38)	73.7%	0.56	A Certified Dementia Practitioner accepted ~79% of the AI’s triage recommendations, supporting the feasibility of telephone-based AI cognitive screening paired with clinician oversight. 6CIT scoring trends toward over-referral and is a priority for model refinement. Limitations. Single site (n = 68); a single, unblinded reviewer (possible anchoring); non-randomized deployment modes. Larger, multi-site, reviewer-blinded studies are warranted.	- Bradford A, et al. Alzheimer Dis Assoc Disord. 2009;23(4):306-314. - Alzheimer’s disease facts and figures. Alzheimers Dement. 2024;20(5):3708-3821. - JaKa MM, et al. Fam Pract. 2024;41(5):761-769. - Brooke P, Bullock R. Int J Geriatr Psychiatry. 1999;14(11):936-940.
Instrument	ICC (95% CI)	Diff*																																
6CIT	0.84 (0.76–0.91)	–1.06																																
Supp. cognitive	0.85 (0.75–0.92)	–0.22																																
IADL	0.93 (0.72–1.00)	+0.10																																
GAD-2	0.93 (0.82–1.00)	0.00																																
GDS-4	1.00 (1.00–1.00)	0.00																																
Deployment Mode	Triage accepted (%)	κ																																
Overall (n=68)	79.4%	0.72																																
Sched. outbound (n=30)	86.7%	0.83																																
In-clinic (n=38)	73.7%	0.56																																
CONTACT																																		
 																																		
 ALZHEIMER'S ASSOCIATION AAIC 26																																		