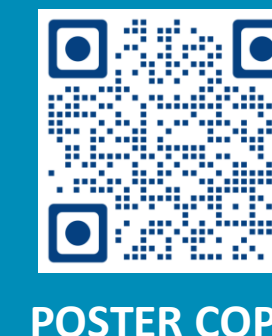


FROM FEASIBILITY TO REAL-WORLD READINESS: INTERPRETING EVIDENCE FOR SELF-ADMINISTERED DIGITAL COGNITIVE ASSESSMENTS

Bryan R. Cobb, PhD, FACMG¹; Miron Derchansky²; Bruno B. Montezano²

1. Eisai, Inc., Nutley, NJ, United States. 2. Blackletter Consulting, Richmond Hill, ON, Canada



POSTER COPY

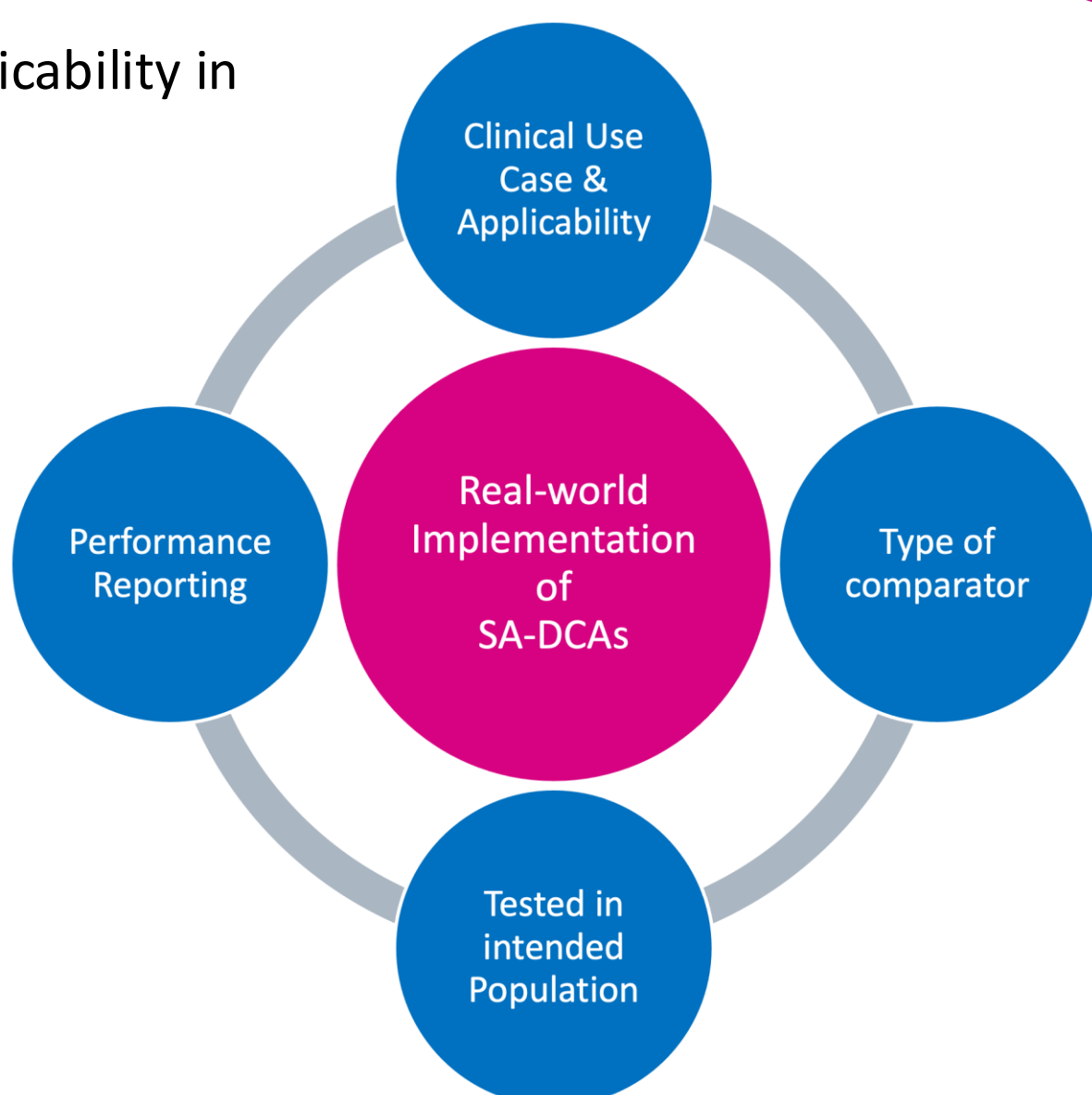
Introduction

- Mild Cognitive Impairment (MCI) remains undiagnosed in >90% of individuals,¹ delaying access to evaluation and treatment, particularly with anti-amyloid therapies in those with MCI due to Alzheimer's Disease (AD)
- Traditional cognitive assessments require trained administrators,² lack sensitivity for early-stage disease leading to diagnostic inaccuracy³ and are difficult to scale in primary care and community settings due to resource constraints⁴
- Remote digital cognitive assessments (DCAs) have the potential to improve screening and detection of cognitive impairment in real-world clinical settings
- Here, we completed an evaluation of the level of evidence in peer-reviewed publications, finding fragmented study designs, disparate results reported in studies that used pre-selected populations with limited biological amyloid confirmation testing
- We introduced key requirements necessary to evaluate self-administered digital cognitive assessments (SA-DCAs), addressing key gaps in validation studies and enabling translation into clinical practice. SA-DCAs offer:
 - Remote, scalable screening (5–30 minutes)
 - Reduced clinician burden with self-testing
 - Potential to extend screening beyond clinic settings with longitudinal tracking
- However, across studies and settings, inconsistent validation approaches make it difficult to interpret how they perform and how should their performance inform real-world screening and downstream diagnostic decisions or whether they are ready for real-world clinical practice
- We sought to introduce a use-case-specific evaluation that could be used to evaluate SA-DCA evidence, namely one that:
 - Enables standardized interpretation of SA-DCA evidence
 - Assesses readiness for real-world MCI screening
 - Aligns digital screening tools with clinical and biomarker-based diagnostic pathways to improve the AD diagnostic and treatment pathways

Methods

- We completed a structured scoping review and synthesis of published SA-DCA validation studies (2012–2025). To do this, we landscaped DCAs that were used remotely, and that could be flexibly deployed (no hardware)
- DCAs were included if they met all the following criteria:
 - Evaluated a SA-DCA that can be remotely deployed and has algorithmic or automated scoring
 - Published >1 manuscript in peer-reviewed journals
 - Reported diagnostic performance metrics
 - Assessed performance in populations reflecting primary-care or community screening contexts “unfiltered populations” (e.g., not pre-selected based on clinical diagnosis or biomarker status)

Figure 1: Assessing Real-World Applicability in Evaluating SA-DCAs



- In parallel, we assessed 4 key domains critical for real-world deployment (**Figure 1**)
 - Population representativeness,
 - Benchmark comparisons,
 - Standardized performance reporting,
 - Clinical use cases and applicability

Methods (continued)

- Electronic/digital versions of existing paper-and-pencil tests were excluded
- For performance reporting, we evaluated whether the following performance characteristics were reported (accuracy, sensitivity, positive and negative predictive values, Cohens D coefficient, Receiver Operator Curves)
- Key study considerations were evaluated:
 - Unfiltered populations
 - Prevalence-adjusted metrics
 - Clinical decision relevance
- We conducted a structured multi-source search from January 2012 to November 2025, key terms and 145 Boolean operators were applied (*red text*; **Figure 2**)
 - The searches covered scientific databases (PubMed and ClinicalTrials.gov), technology and innovation sources: (Google Scholar, Google Patents, Crunchbase, Digital.Health, CBInsights, Growjo, and LinkedIn)
 - Two sequential searches were conducted, tool identification to map all SA-DCAs and peer reviewed publication evidence retrieval (using the same search criteria along with the SA-DCAs identified in the first search)
 - Results are show for final SA-DCAs evaluated and studies filtered.

Figure 2: Search Methodology and Findings

(“Screening” OR “Diagnosis” OR “Identification” OR “Validation” OR “Validity”) AND (“Cognitive Dysfunction” OR “Mild Cognitive Impairment” OR “Cognitive Impairment” OR “Mild Neurocognitive Disorder” OR “Early MCI” OR “MCI” OR “MCI due to AD” OR “Early AD” OR “Cognitive Decline”) AND (“Digital” OR “Wearable” OR “Voice” OR “Vocal” OR “Actigraph” OR “Digital Cognitive Test” OR “Remote” OR “Self-administered” OR “Smartphone” OR “App” OR “Digital Biomarker” OR “Digital Test” OR “Smart Device” OR “Remote Assessment” OR “Remote Administration” OR “At-Home”)

81 SA-DCAs Identified

13 Tools Excluded (not Self-administered)

39 SA-DCAs (>1 study)

SA-DCAs with only a single published study were removed

569 Studies Total

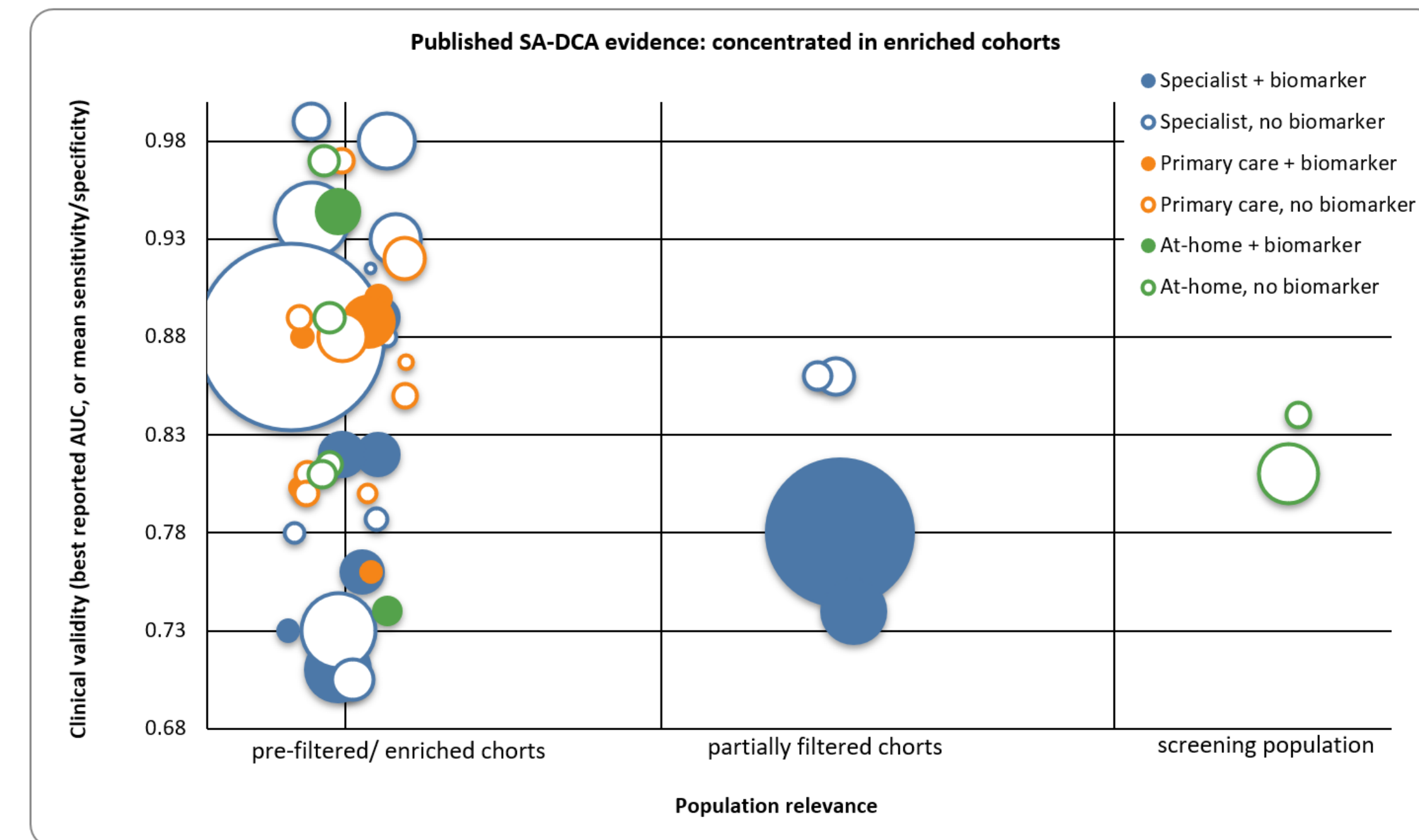
559 studies removed based on exclusion criteria

10 Studies (right population studied) Across 5 SA-DCAs

Results

- Across the literature, we found that reported diagnostic performance frequently reflects pre-selected cohorts (only nine studies were “unfiltered”), unclear performance characteristics, fragmented reporting (n=66 studies didn't report accuracy), and cross-sectional designs that collectively limit generalizability (**Figure 3**)

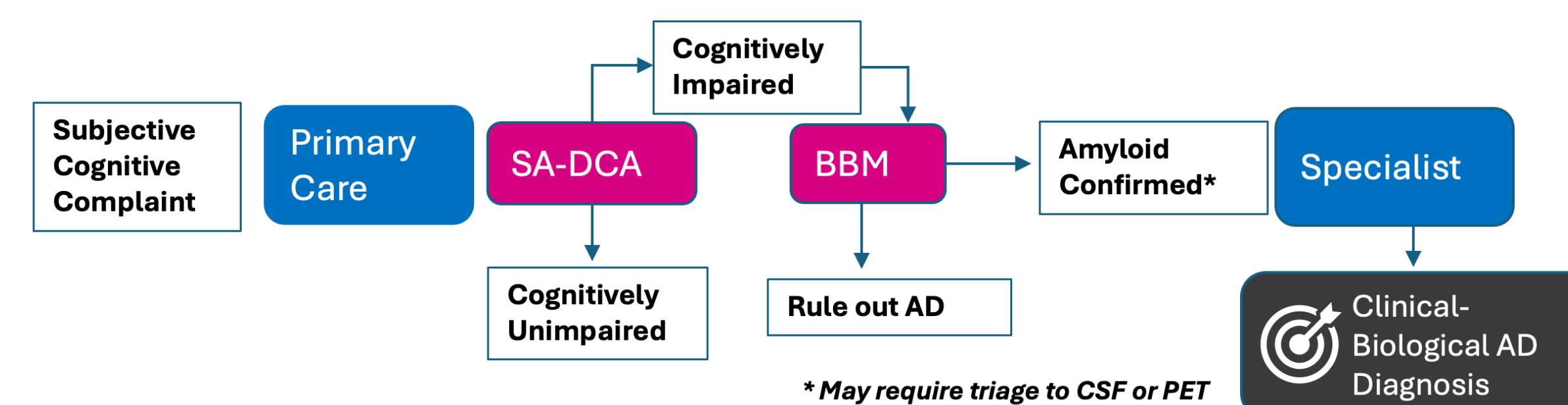
Figure 3: Studies compared based on setting and biomarker confirmatory testing across how the population was selected



Bubble plot of published SA-DCA studies with extractable diagnostic performance data. Population relevance was derived from whether the study sample was pre-filtered/enriched, partially filtered, or aligned with screening-relevant populations. Clinical validity reflects the best reported AUC; when AUC was unavailable, the mean of parsed sensitivity and specificity was used. Bubble size reflects parsed total sample size. Series labels combine study setting (specialist, primary care, or at-home) with whether the study included biomarker anchoring (blood-based biomarker, imaging, or CSF).

- Longitudinal signal stability and clear downstream clinical actions such blood-based biomarker (BBM) confirmatory testing and specialist referral were infrequently evaluated. Clinical-biological characterization is critical to improve the AD diagnostic pathway (**Figure 4**)

Figure 4: Diagnostic Pathway in Alzheimer's Disease: Integrated DCA and BBMs



- Studies often compared preselected cognitively unimpaired populations versus those with combined MCI and moderate dementia inflating AUC estimations
- A key limitation of many digital cognitive assessments not evaluated in this study but should be considered is the lack of large, population-representative normative datasets that adequately account for age, education, sex, language, culture, digital literacy, and device effects

Conclusions

- Current evidence shows heterogeneous validation with few tools meeting minimal clinical readiness criteria
- This study outlines key considerations for moving SA-DCAs from feasibility to real-world readiness
- Using a framework that considers the intended population, rigor of cognitive test comparators, biological confirmation where available, and performance reporting aligned to the intended use case can provide a more transparent and objective assessment of readiness for real-world implementation.
- By establishing a common language for interpreting evidence across digital cognitive screening applications, such a framework can support more consistent evaluation and integration within blood-based biomarker-guided diagnostic pathways and evolving clinical workflow

Acknowledgments/Disclosures

Funding for the analyses and editorial support (Mayville Medical Communications) was provided by Eisai Inc. and Biogen.

References

1. Mattke S, et al. *Alzheimers Res Ther.* 2023;15:128. 2. Livingston G, et al. *Lancet.* 2024;404:572-628. 3. Palmqvist S, et al. *JAMA.* 2024;332:1245-1257. 4. Hampel H, et al. *Neuron.* 2023;111:2781-2799.