

Current State of Self-Administered Brief Computerized Cognitive Assessments for Detection of Cognitive Disorders in Older Adults: A Systematic Review

E. Tsoy¹, S. Zygouris^{2,3,4}, K.L. Possin^{1,4}

1. Department of Neurology, Memory and Aging Center, University of California San Francisco, San Francisco, California, USA; 2. School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece; 3. Network Aging Research, Heidelberg University, Heidelberg, Germany; 4. Global Brain Health Institute, University of California San Francisco, San Francisco, California, USA; Trinity College Dublin, The University of Dublin, Dublin 2, Ireland

Corresponding Author: Katherine L. Possin, PhD, Associate Professor in Residence, Department of Neurology, University of California San Francisco, Memory and Aging Center, Box 1207, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94158, Tel: 415-476-1889, E-mail: katherine.possin@ucsf.edu

Abstract

Early diagnosis of cognitive disorders in older adults is a major healthcare priority with benefits to patients, families, and health systems. Rapid advances in digital technology offer potential for developing innovative diagnostic pathways to support early diagnosis. Brief self-administered computerized cognitive tools in particular hold promise for clinical implementation by minimizing demands on staff time. In this study, we conducted a systematic review of self-administered computerized cognitive assessment measures designed for the detection of cognitive impairment in older adults. Studies were identified via a systematic search of published peer-reviewed literature across major scientific databases. All studies reporting on psychometric validation of brief (≤ 30 minutes) self-administered computerized measures for detection of MCI and all-cause dementia in older adults were included. Seventeen studies reporting on 10 cognitive tools met inclusion criteria and were subjected to systematic review. There was substantial variability in characteristics of validation samples and reliability and validity estimates. Only 2 measures evaluated feasibility and usability in the intended clinical settings. Similar to past reviews, we found variability across measures with regard to psychometric rigor and potential for widescale applicability in clinical settings. Despite the promise that self-administered cognitive tests hold for clinical implementation, important gaps in scientific rigor in development, validation, and feasibility studies of these measures remain. Developments in technology and biomarker studies provide potential avenues for future directions on the use of digital technology in clinical care.

Key words: Computerized cognitive assessment, early detection, mild cognitive impairment, dementia, psychometrics.

Introduction

Dementia remains a widely underdiagnosed condition, both in Western countries (1-5) and globally (6). In light of projected increases in prevalence and burden (7) of dementia, innovative solutions in diagnosis and clinical care of dementia will be critical to alleviate the impact of these changes on public healthcare systems. Most experts agree

that underdiagnosed dementia is a major gap in care because early detection of cognitive decline in older adults with cognitive symptoms (i.e., patient's concerns, informant concerns, etc.) is beneficial for both patients and their caregivers (8-11). Targeted evaluation of cognitive impairment can facilitate early detection of cognitive disorders, which in turn can promote patient safety and wellbeing through more informed medication management, implementation of comprehensive care plans, introduction of lifestyle modifications, improved management of symptoms, and the opportunity to participate in clinical trials (8-11). Additionally, earlier diagnosis of dementia may reduce healthcare costs by decreasing long-term care expenditures with projected economic benefits to the affected individuals and to the public health care systems (12-14).

Primary care providers are in a position to first detect cognitive decline because of their established relationships with their patients. In the United States, cognitive impairment detection in primary care is supported by the Medicare Annual Wellness Visit and a new billing code for cognitive assessment and care planning (11). However, based on a survey of 1,000 PCPs and 1,954 older adults conducted by the Alzheimer's Association, nearly all PCPs (94%) recognized that routine cognitive assessments were important but only 16% of the older adults said that they received routine cognitive assessments (7). Moreover, the survey found that when cognitive assessments are performed in primary care, paper and pencil assessments are almost always used (7). Thus, the most commonly used instruments for detection of cognitive impairment by the PCPs were the Mini Mental State Examination (MMSE; 80%), the Clock Drawing Test (64%), and the Mini-Cog (52%) (7). These conventional paper-based evaluation tools are accurate at detecting dementia but have poorer sensitivity to milder forms of cognitive impairment (15). An additional and important limitation of these conventional tools is that administration, scoring, interpretation, and documentation require substantial clinical staff time. Indeed, among the commonly reported reasons to not

conduct the screening were insufficient time during visits and lack of confidence in performing the evaluation (7, 16).

Computerized tools present both exciting potential advantages and significant challenges for improving the detection of cognitive impairment. Older adults endorse both eagerness to use technology (17) and actual usage of technology in their daily lives (18), particularly touchscreen devices, which allow for direct interaction and have lower motor demands and relative ease of use (17, 19). Technology-based assessments offer enhanced precision of measurement and scoring, instant automated scoring and interpretation, standardized administration, enhanced stimulus presentation, availability of multiple alternate forms to minimize practice effects, and potential for adaptive capabilities and more sophisticated algorithms (20-23). Additionally, computerized tools have been discussed as being more cost-effective, particularly with respect to materials and supplies.

Computerized measures also present with challenges related to examinee variables (familiarity with technology, attitude and anxiety towards technology) (24) and technological issues (variability in hardware and software characteristics, data and privacy issues, data charges and internet access) (20, 22, 23). Additionally, a number of past reviews highlighted the lack of adequately established psychometric standards, limited or unfamiliar response modality, and poorly designed user interface (22, 25). A number of studies also demonstrated that computerized measures failed to demonstrate equivalence between the examinee's experience of computer versus traditional test administration (e.g., participants performed worse on electronic version of the Montreal Cognitive Assessment [MoCA] compared to paper MoCA) (26).

Although most brief cognitive assessments designed for primary care use are examiner-administered, self-administered instruments have the potential to minimize costs and practice barriers related to training and staffing costs (15, 27), and to support social distancing requirements during the COVID-19 pandemic. Patients could complete a self-administered brief cognitive assessment at the clinic prior to or following a provider appointment, or at home on their own device. Additional benefits include minimization of examiner effects and greater accessibility of the tool for patients in remote locations (23). If the self-administered test is translated into different languages, accessibility may be increased for patients who do not speak English because interpreter may not be required. Finally, there is some evidence that absence of an examiner may reduce observer-related stress and increase respondents' openness during administration, although these findings were reported only in studies on unsupervised surveys (27). While self-administered assessments share the same challenges with examiner-administered computerized assessments, additional challenges include lack of monitoring to ensure response validity (compliance, effort, motivation), lack

of support should the patient need help with a task or runs into technological issues, and loss of qualitative data available from a conventional in-person evaluation (25). Additionally, a bring-your-own-device paradigm may present additional challenges related to potential technological differences to ensure consistent stimuli presentation and reaction time measurement, such as screen size and resolution, operating system, central processing unit capacity, etc. (23).

In this study, we conducted a systematic review of studies on self-administered computerized assessments designed to detect mild cognitive impairment (MCI) and dementia in older adults, discuss benefits and weaknesses related to their use, and provide practical recommendations and considerations regarding implementation of these measures into clinical practice.

Method

Databases

A systematic search of published literature was conducted from February 1, 2020 to April 20, 2020. Databases searched included PubMed, EMBASE, Web of Science, and PsycINFO. Example search items used were "computerized" or "tablet," and "cognitive assessment" or "cognitive screen," and "older adult" or "geriatric," and "dementia" or "cognitive impairment." Specific search strings for each database are included in Supplementary Methods. Additional search filters included 1) original peer-reviewed articles (not book chapters, abstracts or conference papers, unpublished dissertations, or review studies), 2) studies published on or after January 1, 2000, and 3) studies published in English language.

Inclusion and exclusion criteria

Inclusion criteria were: 1) studies including a control sample and a clinical (MCI or all-cause dementia diagnosed based on published consensus criteria; 28-34) sample of older adults (ages 50 years and above); 2) studies reporting on a brief (administration time of 30 minutes or less) computerized cognitive assessment tool; and 3) studies reporting on psychometric characteristics of the measure, including reliability and validity indices.

Exclusion criteria were: 1) studies reporting on the use of the cognitive assessment tool in individuals with medical conditions other than MCI or dementia (e.g., schizophrenia, multiple sclerosis, etc.); 2) studies reporting on computerized instruments that were not cognitive assessment tools (e.g., informant surveys, functional questionnaires, etc.); 3) studies reporting on computerized instruments that required a dedicated hardware platform for use (e.g., virtual reality sets, hardware kiosks, etc.) due to potential barriers of implementing these modalities in clinical settings; 4)

Table 1. Quality assessment scale

Domains: assessment of core cognitive domains, including attention/executive functions, memory, language, and visuospatial functions
1 = Narrow focus (2 domains)
2 = Moderate coverage of domains (3 domains)
3 = Comprehensive coverage of domains (≥ 4 or more domains)
Validation sample: sample size and use of standard diagnostic criteria
1 = Low sample size (<50 per group) and no reference to standard diagnostic criteria
2 = Low sample size (<50 per group) or no reference to standard diagnostic criteria
3 = Adequate sample size (≥ 50 per group) and reference to standard diagnostic criteria
Reliability: internal consistency and test-retest stability
1 = No data
2 = 1 type of reliability
3 = ≥ 1 type of reliability
Validity: concurrent validity and criterion validity
1 = No data
2 = 1 type of validity
3 = ≥ 1 type of validity
Level of examiner involvement
1 = High: examiner present and provides frequent assistance with administration or scoring requires an examiner
2 = Moderate: examiner present and provides minimal assistance and scoring is automated
3 = Low: examiner not present and scoring is automated
Availability for clinical use
1 = Available but not usable for clinical application or no data on availability
2 = Available for clinical use and requires purchase of a dedicated device
3 = Available for clinical use and does not require purchase of a dedicated device
Availability in different languages
1 = Available in 1 language or no data on availability of different languages provided
2 = Available in 2 languages
3 = Available in ≥ 3 languages
Feasibility studies
1 = No feasibility studies conducted in intended use settings
3 = Feasibility studies conducted in intended use settings
Data security: encryption, compliance with healthcare regulations
1 = Data security not addressed or no data on security provided
2 = Data security addressed, but information provided is limited
3 = Comprehensive description of data security and compliance with regulations
Delivery of test results
1 = Only numerical results reported or no data on results reporting provided
2 = Interpretation of numerical results reported (not immediate) and no/unclear information on care guidance
3 = Interpretation of numerical results reported (immediate) and care guidance provided

studies reporting on computerized cognitive measures that were not validated in English; 5) studies reporting on computerized cognitive measures that assess a single cognitive domain; and 6) studies reporting on instruments that required an examiner to administer the tests.

Data extraction

To minimize selection bias, two authors (ET and SZ) independently conducted database searches using same search terms and reviewed titles and abstracts for inclusion criteria. Same two authors then conducted a full-text review of screened-in studies for exclusion criteria. Data extracted from the final set of studies

included in the review were: 1) platform; 2) level of supervision required; 3) administration time; 4) characteristics of the validation samples; and 5) psychometric indices. We extracted additional data on commercial availability of the tool, requirements for devices, automated reporting of results, available languages, and number of associated publications based on review of bibliography and dedicated websites of the measures when available. A brief e-mail survey (Supplementary Methods) was also sent out to test developers to collect additional information. Any disagreements were resolved via consensus discussions with a third reviewer (KLP).

Quality assessment

The quality of the studies was assessed using a scale specifically designed for this study in order to capture important considerations and factors for self-administered cognitive assessments (Table 1). Development of the criteria included in the scale was based on prior works on computerized cognitive tools (20). Specifically, we assessed the measures based on the following criteria: comprehensive assessment of the core cognitive domains, size of the validation sample and use of standard diagnostic criteria for identifying participants with cognitive impairment, reliability and validity indices, degree to which an examiner is involved in the testing process in relation to fully automated procedures, current availability for clinical use including any requirements for purchase of a dedicated device, availability of offered tests in multiple languages for participants/patients whose first language is other than English, presence or absence of feasibility studies in the intended settings, issues related to data security and compliance with regulations, and comprehensiveness of the delivery of test results (Table 1). Any ambiguity or disagreements were resolved via a consensus agreement with the third author (KLP).

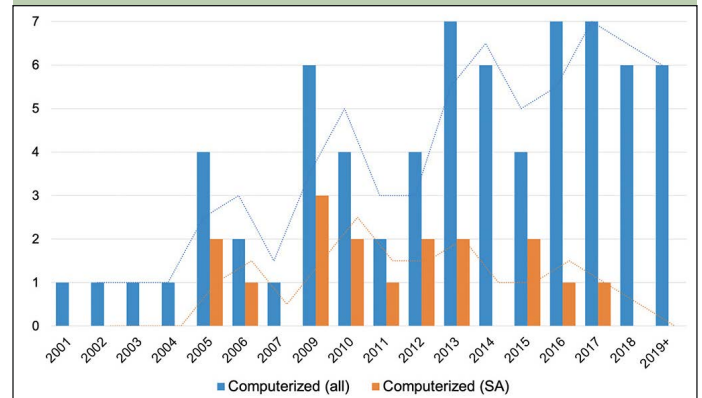
Results

The search identified 11,617 citations which, after removal of duplicates, resulted in 9,986 unique records. Seventy studies, which were selected from the initial screening process, were further assessed for eligibility based on full-text review. The overall trends in peer-reviewed published studies on computerized tools which were included in the full-text review are presented in Figure 1. In total, 17 studies reporting on 10 self-administered computerized tools were included in the review (for PRISMA flowchart, see Supplementary Results).

The measures included in this review were Computer Assessment of Memory and Cognitive Impairment (CAMCI; 35), Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment (CANS-MCI; 36, 37), Computerized

Cognitive Screening (CCS; 38), CNS Vital Signs (CNSVS; 39,40), Computerized Self Test (COGSelfTest; 41), CogState (CogState; 42-44), CogState Brief Battery (CogState BB; 45-48), Cognitive Testing on Computer (C-TOC; 49), digitally translated Self-Administered Gerocognitive Examination (eSAGE; 50), and an untitled test developed by Kluger et al. (51). Results of the quality assessment ratings of the included tools are reported in Table 2.

Figure 1. Number of published peer-review studies included in the full-text review from January 1, 2000 to April 20, 2020



Abbreviations: SA, self-administered. Blue bars represent the number of studies that met inclusion criteria but were not self-administered, while orange bars represent the number of studies that met inclusion criteria and were self-administered.

Tool characteristics

Detailed characteristics of the measures included in the review are reported in Table 3. Of the 10 tools included in the review, 3 were available only on a personal computer (PC) platform (CNSVS, COSSelfTest, untitled test), 2 only on a tablet (CAMCI, CCS), and 4 on both platforms (CANS-MCI, CogState, CogState BB, C-TOC, eSAGE). CogState measures were available on an unrestricted set of devices, while most other tools had some restrictions, such as requiring touchscreen capabilities (CANS-MCI, CCS, eSAGE), keyboard input (CNSVS, COGSelfTest, untitled test), or a specific set of devices (CAMCI, C-TOC). Regarding level of supervision, CAMCI and CANS-MCI were designed to be administered in medical and research settings and test developers do not recommend at-home remote testing. Additionally, C-TOC and eSAGE require a trained examiner for scoring. Administration times varied across measures averaging at approximately 15-20 minutes across tools. Commercial availability was a common characteristic with 8/10 measures, except for C-TOC and untitled test, available for purchase. Finally, more than half (6/10) of the measures had fewer than 5 peer-reviewed published studies on the use of the measure across any age groups or clinical populations, while the remaining 4 tools (CAMCI, CNSVS, CogState, CogState BB) were researched more widely with at least 10 peer-reviewed published studies.

Table 2. Quality assessment ratings of included measures

	Domains	Validation sample	Reliability	Validity	Examiner involvement	Clinical availability	Languages	Feasibility studies	Data security	Result delivery	Mean score
CAMCI	2	2	2	2	3	3	1	3	2	2	2.2
CANS-MCI	2	2	3	3	3	3	3	1	3	3	2.6
CCS	2	2	2	3	2	1	1	1	1	1	1.6
CNSVS	1	1	2	3	3	3	3	1	3	2	2.2
COGSelfTest	3	2	3	3	3	2	1	1	1	1	2.0
CogState	1	1	2	3	2	2	3	1	3	2	2.0
CogState BB	1	3	2	3	2	2	3	3	3	2	2.4
C-TOC	3	2	1	3	1	1	1	1	1	1	1.5
eSAGE	3	2	1	3	1	3	2	1	2	2	2.0
Untitled test	2	1	1	3	2	1	1	1	1	1	1.4

Abbreviations: CAMCI, Computer Assessment of Memory and Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CCS, Computerized Cognitive Screening; CNSVS, CNS Vital Signs; COGSelfTest, Computerized Self Test; CogState BB, CogState Brief Battery; C-TOC, Cognitive Testing on Computer; eSAGE, digitally translated Self-Administered; Gerocognitive Examination.

Validation samples

Characteristics of the validation samples are presented in Table 4. There was substantial variability in sample sizes across studies, and only 2/10 instruments (CAMCI, CogState BB) were validated in large cohorts with at least 50 or more participants in each diagnostic group. Studies on more than half of the instruments (6/10) used published criteria to classify participants into diagnostic groups, while 3 measures were validated in samples classified by scores on standard cognitive testing (CAMCI, CNSVS, CogState), and the study on the untitled test did not provide sufficient details regarding diagnostic criteria used. Additionally, studies were varied with regard to selection of cognitively normal participants ranging from recruitment of spouses of individuals with MCI and dementia as controls to comprehensive assessment of control subjects. With regard to demographic characteristics of validation samples, the vast majority of participants were non-Hispanic White with 12 or more years of education (educational attainment was not reported for CNSVS and untitled test).

Psychometric properties

Test-retest reliability was reported on 6/10 tools (Table 4), and there was substantial variability in the reported indices both across and within individual measures by constituent subtests. Reported reliability coefficients of 5/10 measures (CANS-MCI, CNSVS, COGSelfTest, CogState, CogState BB) were consistently within ranges of moderate to high stability based on standard psychometric criteria (52). The range of time intervals for test-retest reliability studies also varied ranging from 2 hours to 12 months. Internal consistency estimates were reported on 3/10 tools (CANS-MCI, CCS, COGSelfTest) with coefficients ranging from .43-.97.

Table 4 presents concurrent validity estimates with either paper-and-pencil brief cognitive assessments or conventional neuropsychological tests which differed by site and study design for 9/10 measures (not reported for CAMCI). Similar to reliability findings, these indices varied significantly across and within measures with most tools demonstrating moderate degree of concurrent validity with standard tests. Across studies, highest concurrent validity estimates were mostly reported with standard brief global assessments (e.g., MMSE), while domain-specific concurrent estimates were more varied. Criterion validity estimates were reported for all measures including discriminant analyses results for 9/10 tools and mean group differences on 1 tool (Table 4). Studies on 6/10 instruments (CAMCI, CANS-MCI, CNSVS, CogState, CogState BB, eSAGE) reported on discriminant analyses between control and MCI groups with sensitivity indices ranging from .41 to .90 and corresponding specificity estimates of .64-.94. As expected, the indices for distinguishing cognitively normal and dementia groups reported on 3/10 tools (CCS, CNSVS, CogState BB) were slightly higher with a sensitivity range of .53-.94 and a specificity range of .50-.94. Finally, discrimination between cognitively normal and impaired (MCI and dementia combined) groups was reported for 3/10 instruments (COGSelfTest, eSAGE, untitled test) with the sensitivity range of .71-.99 and the specificity range of .72-.95. The remaining study on C-TOC reported on criterion validity in the form of mean group differences suggesting significantly lower performance on tasks of episodic memory, executive functions, and speed measures in cognitively impaired no dementia group compared to controls.

Table 3. Summary of the features of included measures

Test	Platform	Device requirements	Time (min)	Available commercially	Automated report	Available languages (N)	Publication(s) included in review	Associated publications (N)*
CAMCI	Tablet	Surface Pro 3/4	20-25	Yes	Yes	1	Saxton 2009	12
CANS-MCI	PC/Tablet	Touchscreen	30	Yes	Yes	4	Ahmed 2012, Tornatore 2005	3
CCS	Tablet	Touchscreen	3	Yes	ND	ND	Scanlon 2016	1
CNSVS	PC	Non-touchscreen	30	Yes	Yes	60	Gualtieri 2005, Gualtieri 2006	100+
COGSelfTest	PC	Non-touchscreen	6-15	Yes	ND	ND	Dougherty 2010	1
CogState	PC/Tablet	Any device	20-30	Yes	Yes	43	de Jager 2009, Hammers 2011, Lim 2013	100+
CogState BB	PC/Tablet	Any device	12-15	Yes	Yes	43	Hammers 2012, Fredrickson 2010, Maruff 2013, Mielke 2015	100+
C-TOC	PC/Tablet	iPad	30-45	No	ND	1	Jacova 2015	1
eSAGE	PC/Tablet	Touchscreen	14-17	Yes	Yes	2	Scharre 2017	1
Untitled test	PC	Non-touchscreen	12-15	No	ND	1	Kluger 2009	1

Abbreviations: CAMCI, Computer Assessment of Memory and Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CCS, Computerized Cognitive Screening; CNSVS, CNS Vital Signs; COGSelfTest, Computerized Self Test; CogState BB, CogState Brief Battery; C-TOC, Cognitive Testing on Computer; eSAGE, digitally translated Self-Administered Gerocognitive Examination; ND, no data; PC, personal computer; *Includes peer-reviewed published journal articles across all age groups and clinical populations.

Delivery of results and available languages

Automated delivery of results is a feature of 6/10 tools (CAMCI, CANS-MCI, CNSVS, CogState, CogState BB, eSAGE), all of which include interpretation of results based on normative adjustments and differ in presentation of the results (Table 3). Automated reports on 2 measures (eSAGE, CANS-MCI) are reviewed by a trained professional prior to generation and the delivery of the results is thus not immediate. Additionally, CANS-MCI reports include recommendations to physicians on next steps and potential areas for intervention, and eSAGE has functionality to generate patient-facing reports. User-friendliness of the automated reports for non-specialist physicians was empirically examined only for CAMCI (53).

Data on available languages was reported on 6/10 tools (Table 3). Out of these, CAMCI, C-TOC, and untitled test are available only in English, eSAGE is available in English and Spanish, CANS-MCI is available in 4 languages, and CNSVS, CogState, and CogState BB are available in more than 40 languages. Additionally, availability of peer-reviewed studies on validation of non-English versions of these tools for detection of cognitive disorders in older adults varied substantially with the majority of published studies conducted in English-speaking samples.

Discussion

In this systematic review, we evaluated 10 brief self-administered computerized cognitive assessment measures designed to detect cognitive disorders in older adults. Similar to past reviews of computerized cognitive tools (20, 21), we found significant variability across measures with regard to characteristics and design of the tools, sizes of validation samples, availability in

different languages, and psychometric qualities, all of which are crucial considerations for potential widescale implementation of these measures in clinical care. Specifically, we found that few of the reviewed measures were validated in sufficiently large samples (CAMCI, CogState BB) and are available in multiple languages (CANS-MCI, CNSVS, CogState, CogState BB). Test-retest reliability, which is critical for self-administered tools aiming to monitor cognitive functions, was reported only on 60% of the tools, and internal consistency measures were reported on even fewer measures. While almost all reviewed measures reported data on concurrent validity, the estimates for several individual domain subtests within some tools were low. These findings are concerning, particularly when considering the need for a battery to distinguish among different types of MCI and dementia and inform differential diagnoses in non-specialty settings (16). On the other hand, we found that most measures required minimal involvement of an examiner in test administration and scoring of results and were available as standalone applications on several device types (e.g., PC, tablet computer, etc.). These features are important benefits of self-administered computerized tools, particularly if additional built-in functionality for integration of results into electronic medical records (EMR) systems is developed (54). In general, despite the promise that self-administered cognitive tests hold for clinical applications, important gaps in scientific rigor in development, validation, and feasibility studies of these measures remain. Below we discuss critical areas of need for future development and validation of self-administered cognitive measures that would facilitate their potential for widescale clinical implementation.

One of the most critical gaps identified in the current review is the size and demographic constitution of the validation samples. In particular, several studies included

Table 4. Summary of the psychometric properties of included measures

	Domains				Validation Sample	Diagnostic Criteria	Reliability			Validity
	Attention / EF	Memory	Language	VS			Internal consistency	Test-Retest	Concurrent	
CAMCI	X	X		X	N = 524 296 CN: age 72±6, education 14±3 228 MCI: age 75±7, education 13±3	Based on standard cognitive tests	ND	3 w r = .30-.74	ND	MCI: SN = .86, SP = .94
CANS-MCI	X	X	X		N = 35 20 CN: age 77±4, education 15±3 15 MCI: age 81±7, education 13±3	(29)	$\alpha = .77-.96$	1 m r = .61-.85 6 m r = .62-.89	r = .44-.64	MCI: SN = .89, SP = .73
CCS	X	X		X	N = 60 20 CN: age 73±12, education 55% 12+ 40 DEM: age 77±13, education 50% 12+	(32)	$\alpha = .79^*$	ND	r = .38-.78	DEM: SN = .94, SP = .60
CNSVS	X	X			N = 178 89 CN: age 63, education ND 36 MCI: age 66, education ND 53 DEM: age 62, education ND	Based on standard cognitive tests	ND	4 w r = .65-.88	r = .26-.79	MCI: SN = .90, SP = .64-.85 DEM: SN = .90, SP = .50-.94
COGSelfTest	X	X	X	X	N = 215 104 CN: age 75±7, education 15±3 27 MCI: age 67±6, education 14±4 84 AD DEM: age 75-77, education 13-15	(28,33)	$\alpha = .43-.88$	6 w r = .76	r = .56	MCI+DEM: SN = .99, SP = .95
CogState	X	X			N = 119 98 CN: age 77±6, education mostly 12+ 21 MCI: age 82±5, education mostly 12+	Based on standard cognitive tests	ND	2 h r = .54-.80 3 m r = .76-.97	r = .20-.53	MCI: SN = .78, SP = .90
CogState BB	X	X			N = 817 659 CN: age 70±7, education median 12 107 MCI: age 76±8, education median 12 51 AD DEM: age 79±7, education median 12	(28,31,33)	ND	4 m r = .90-.96 12 m r = .65-.91	r = .45-.52	MCI: SN = .41-.80, SP = .85-.86 DEM: SN = .53-1.0, SP = .85-.86
C-TOC	X	X	X	X	N = 49 16 CN: age 68±8, education 94% 16+ 16 CIND: age 64±7, education 81% 16+	(32,33)	ND	ND	r = .37-.88	CIND: significant mean differences on memory, EF, and speed tests
eSAGE	X	X	X	X	N = 66 21 CN, 24 MCI, 21 DEM, age (all) 75±7, education (all) 15±3	(28,33)	ND	ND	r = .67-.76	MCI: SN = .63, SP = .81 MCI+DEM: SN = .71, SP = .90
Untitled test	X	X	X		N = 105 39 CN: age 64±11, education ND 19 MCI: age 72±10, education ND 47 DEM: age 69-78, education ND	No reference to consensus criteria	ND	ND	r = .49-.80	MCI+DEM: SN = .92, SP = .72

Abbreviations: AD, Alzheimer's disease; CAMCI, Computer Assessment of Memory and Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CCS, Computerized Cognitive Screening; CIND, cognitively impaired no dementia; CN, cognitively normal; CNSVS, CNS Vital Signs; COGSelfTest, Computerized Self Test; CogState BB, CogState Brief Battery; C-TOC, Cognitive Testing on Computer; DEM, dementia; EF, executive functions; eSAGE, digitally translated Self-Administered Gerocognitive Examination; h, hours; m, months; MCI, mild cognitive impairment; ND, no data; PC, personal computer; w, weeks; *Reported only in the dementia group

in this review included fairly small (<50 participants in each diagnostic group) validation samples, and the majority of validation samples were comprised of White, highly educated individuals. Because we did not identify systematic reporting of the power analyses for detecting main effects in the reviewed studies, we applied a generous estimate of 50 participants per group as part of our criteria. However, given the variability in statistical approaches used in these studies, reporting of robust power estimations would not only support the overall results but also ensure transparency, comparability, and generalizability of results across cohorts.

Another important finding of this review is the scarcity of feasibility and implementation studies of self-administered instruments in care settings. In contrast to highly standardized research settings, self-administration of cognitive assessments in the real world may be subject to interruptions and practical limitations such as time and space, which could be detrimental not only to feasibility but also to the validity of results (22). Some domains, such as orientation, may not be applicable for self-administration altogether, as it would be difficult to ensure the fidelity of responses on such tasks in absence of examiner. Given these considerations, research on development and validation of self-administered computerized measures must be supported by well-designed feasibility and implementation studies, which will critically inform the clinical utility of these measures in intended settings. Specifically, feasibility and implementation studies have the potential to identify facilitators and barriers to clinical applications, inform development of optimal diagnostic and care pathways, and, based on the insights from 2 measures (CAMCI, 53, 55 and CogState BB, 56) studied in clinical settings, are critical for informing targeted solutions for individual practices.

The automated delivery of results is key to the clinical utility of computerized tools. To facilitate integration of self-administered tests in non-specialty settings, they should have easy-to-interpret, safe automated report delivery, which would ideally inform the provider on follow-up care and diagnostic considerations based on evidence-based practice guidelines (54). Out of the measures reviewed, only CANS-MCI features an automated report that provides such recommendations to physicians. Moreover, a study on CAMCI with primary care physicians (53) suggested that providers expressed a need for training in interpretation of the report, which highlights the need for refinement of automated reporting and empirical studies on non-specialty providers' attitudes and perceptions of cognitive testing results.

With regard to patient-level characteristics, there are number of critical considerations, particularly given the dearth of normative or validation data in older adults who are racially/ethnically diverse and have low educational attainment. Importantly, one of the prior studies on CogState BB reported that older adults with

lower education were less likely to meet the integrity criteria on 3/5 subtests of the battery (47). This is a major issue given that one of the most promising potentials of self-administered cognitive assessments is supporting services in remote areas and populations less likely to seek specialty evaluations. Moreover, numerous studies suggest that older adults in the U.S. who report Hispanic ethnicity, non-White race, or low education are at a higher risk for neurodegenerative diseases (7) and experience significant disparities in healthcare access and delivery (57). Well-validated self-administered assessments may help substantially reduce these disparities given their potential to deliver tests in different languages (23) but only if they undergo rigorous scientific and cross-cultural validation development. In addition to language and education variables, it is important to validate computerized tools across socioeconomic groups, as past evidence suggests that older adults with lower socioeconomic status reported lower levels of intention to use computerized cognitive testing (24). Finally, successful clinical implementation of even the most well-validated tools would likely require continuous efforts for education and outreach to patients belonging to underrepresented groups as well as their medical providers and families.

Another important variable to consider for self-administration of computerized cognitive measures is the impact of familiarity with technology on test results. While some studies (CCS; 38) reported no differences in test scores between older adults with and without technology experience, these variables do appear to play a significant role through interactions with age (CogState BB; 48) and diagnostic status (C-TOC; 49). Moreover, comparisons between content-equivalent paper and electronic version measures revealed that older adults with no technology experience performed worse on the electronic version of the measure compared to those with digital proficiency (eSAGE; 50). Finally, while most studies examine the associations between familiarity with technology and computerized cognitive testing results on a group-based level, systematic research on the impact of these variables for individual patients is necessary to support utility of self-administered assessments in clinical practice.

Regarding technical considerations, the practice parameters on optimal development and validation of computerized cognitive tools, including issues related to end-user agreements, privacy, data security and reporting (23, 25), are highly relevant to self-administered paradigms. Of particular relevance are challenges related to the use of bring your own device (BYOD) model and dependence on broadband connection, which pose a threat for timing and measurement error and may thus lead to inaccurate interpretation of results. Built-in integrity measures designed to address this challenge are features of some self-administered tools (e.g., CogState; 47, 48), but are not widely available across reviewed

instruments. Moreover, because of rapidly evolving hardware and operating systems across both PC and tablet platforms, computerized assessments require continuous quality assurance testing and software maintenance investments, and these challenges are greater when many devices (i.e., BYOD models) are supported. Additionally, availability of a measure on multiple devices also requires supporting research to establish the equivalence of normative and psychometric data across different platforms and input parameters, such as touchscreen vs. keyboard response, screen size, etc. Finally, past studies highlighted concerns regarding underreporting of privacy and security safeguards and their limitations on currently available computerized measures (22), and test developers should strive to explicitly disclose any potential consequences of data loss or breaches, particularly for individual patients in clinical settings. As such, collaborative efforts among researchers, funding bodies, industry, policy regulators, and consumers are necessary to develop robust, sustainable platforms supporting optimal levels of security, privacy, confidentiality, and potential functionality of data sharing across sites in order to promote and maintain successful implementation of computerized tools into everyday clinical practice while meeting programming cost demands.

This study has a number of limitations. First, while all attempts were made to conduct a comprehensive search of available literature, our results were limited to studies available in databases searched. Second, due to variability in study design and test statistics, quantitative summary of the findings was not possible. Finally, our review was limited to inclusion of studies that reported on instruments available at least in English language, and a number of promising self-administered computerized cognitive measures validated in non-English cohorts were not considered.

At the same time, a major strength of this study is the scope of reviewed characteristics of the included measures, including not only psychometric qualities but also functional and technological features critical for clinical implementation considerations. Additionally, our review is conducted at a point in time when the need for self-administered cognitive assessments has never been so dire in both clinical and research settings. In light of rapidly developing technologies for identifying disease biomarkers, future studies should examine the associations of a variety of self-administered cognitive assessments with biomarkers of neurodegenerative diseases, particularly given promising existing studies within this line of research (CogState; 58 and Computerized Cognitive Composite [C3]; 59). Additionally, future studies on self-administered cognitive measures in clinical settings should explore optimal implementation paradigms and provider behavior patterns which would be valuable for informing public healthcare policy and efforts to support earlier diagnosis of cognitive disorders in older adults.

Conclusion

In conclusion, this systematic review identified 10 self-administered brief computerized cognitive measures which have a potential for future clinical implementation. Continuous collaborative efforts of different stakeholders are necessary to address the gaps in scientific rigor of development, validation, and implementation studies of these measures.

Funding: This study was supported by the National Institute of Neurological Disorders and Stroke [UG3 NS105557] and the Global Brain Health Institute. The sponsors had no role in the design and conduct of the study; in the collection, analysis, and interpretation of data; in the preparation of the manuscript; or in the review or approval of the manuscript.

Declarations of interest: Elena Tsoy declares no conflicts of interest. Stelios Zygouris declares no conflicts of interest. Katherine L. Possin has received research funding from the NIH, Quest Diagnostics, the Global Brain Health Institute, the Merck Foundation, and the Rainwater Charitable Foundation, consulting fees from ClearView Healthcare Partners and Vanguard, and a speaking fee from Swedish Medical Center.

References

- Chodosh J, Petitti DB, Elliott M, et al. Physician recognition of cognitive impairment: evaluating the need for improvement. *J Am Geriatr Soc.* 2004;52(7):1051-1059. doi:10.1111/j.1532-5415.2004.52301.x
- Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Underdiagnosis of dementia in primary care: variations in the observed prevalence and comparisons to the expected prevalence. *Aging Ment Health.* 2011;15(8):978-984. doi:10.1080/13607863.2011.596805
- Savva GM, Arthur A. Who has undiagnosed dementia? A cross-sectional analysis of participants of the Aging, Demographics and Memory Study. *Age Ageing.* 2015;44(4):642-647. doi:10.1093/ageing/afv020
- Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. *Alzheimer Dis Assoc Disord.* 2009;23(4):306-314. doi:10.1097/WAD.0b013e3181a6bebc
- Amjad H, Roth DL, Sheehan OC, Lyketsos CG, Wolff JL, Samus QM. Underdiagnosis of dementia: an observational study of patterns in diagnosis and awareness in US older adults. *J Gen Intern Med.* 2018;33(7):1131-1138. doi:10.1007/s11606-018-4377-y
- Lang L, Clifford A, Wei L, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open.* 2017;7(2):e011146. doi:10.1136/bmjopen-2016-011146
- Alzheimer's Association, 2019. Facts and Figures. <https://www.alz.org/media/documents/alzheimers-facts-and-figures-2019-r.pdf>. Accessed June 30, 2020.
- U.S. Department of Health & Human Services, 2019. National plan to address Alzheimer's disease: 2019 update. <https://aspe.hhs.gov/system/files/pdf/262601/NatlPlan2019.pdf>. Accessed August 12, 2020.
- Schneider J, Jeon S, Gladman JT, Corriveau RA, 2019. ADRD summit 2019 report to the National Advisory Neurological Disorders and Stroke Council. https://www.ninds.nih.gov/sites/default/files/2019_adrd_summit_recommendations_508c.pdf. Accessed August 12, 2020.
- U.S. Centers for Disease and Control Prevention, 2019. Advancing early detection: A healthy brain initiative issue map. <https://www.cdc.gov/aging/healthybrain/pdf/Issue-Map-Early-Detection-508.pdf>. Accessed August 12, 2020.
- Cordell CB, Borson S, Boustani M, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. *Alzheimers Dement.* 2013;9(2):141-150. doi:10.1016/j.jalz.2012.09.011
- Alzheimer's Association, 2018. Facts and Figures. <https://www.alz.org/media/homeoffice/facts%20and%20figures/facts-and-figures.pdf>. Accessed June 30, 2020.
- Barnett JH, Lewis L, Blackwell AD, Taylor M. Early intervention in Alzheimer's disease: a health economic study of the effects of diagnostic timing. *BMC Neurol.* 2014;14:101. doi:10.1186/1471-2377-14-101
- Weimer DL, Sager MA. Early identification and treatment of Alzheimer's disease: social and fiscal outcomes. *Alzheimers Dement.* 2009;5(3):215-226. doi:10.1016/j.jalz.2009.01.028
- Scott J, Mayo AM. Instruments for detection and screening of cognitive

- impairment for older adults in primary care settings: a review. *Geriatr Nurs*. 2018;39(3):323-329. doi:10.1016/j.gerinurse.2017.11.001.
16. Bernstein A, Rogers KM, Possin KL, et al. Dementia assessment and management in primary care settings: a survey of current provider practices in the United States. *BMC Health Serv Res*. 2019;19(1):919. doi:10.1186/s12913-019-4603-2
 17. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. *Front Psychol*. 2017;8:1687. doi:10.3389/fpsyg.2017.01687
 18. Pew Research Center, 2018. Mobile fact sheet. <http://www.pewinternet.org/fact-sheet/mobile/>. Accessed June 30, 2020.
 19. Canini M, Battista P, Della Rosa PA, et al. Computerized neuropsychological assessment in aging: testing efficacy and clinical ecology of different interfaces. *Comput Math Methods Med*. 2014;2014:804723. doi:10.1155/2014/804723
 20. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. *Alzheimers Dement*. 2008;4(6):428-437. doi:10.1016/j.jalz.2008.07.003
 21. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. *Am J Alzheimers Dis Other Dement*. 2015;30(1):13-28. doi:10.1177/1533317514522852
 22. Gates NJ, Kochan NA. Computerized and on-line neuropsychological testing for late-life cognition and neurocognitive disorders: are we there yet? *Curr Opin Psychiatry*. 2015;28(2):165-172. doi:10.1097/YCO.0000000000000141
 23. Parsons TD, McMahan T, Kane R. Practice parameters facilitating adoption of advanced technologies for enhancing neuropsychological assessment paradigms. *Clin Neuropsychol*. 2018;32(1):16-41. doi:10.1080/13854046.2017.1337932
 24. Werner P, Korczyn AD. Willingness to use computerized systems for the diagnosis of dementia: testing a theoretical model in an Israeli sample. *Alzheimer Dis Assoc Disord*. 2012;26(2):171-178. doi:10.1097/WAD.0b013e318222323e
 25. Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Clin Neuropsychol*. 2012;26(2):177-196. doi:10.1080/13854046.2012.663001
 26. Wallace SE, Donoso Brown EV, Simpson RC, et al. A Comparison of electronic and paper versions of the Montreal Cognitive Assessment. *Alzheimer Dis Assoc Disord*. 2019;33(3):272-278. doi:10.1097/WAD.0000000000000333
 27. Feenstra HE, Vermeulen IE, Murre JM, Schagen SB. Online cognition: factors facilitating reliable online neuropsychological test results. *Clin Neuropsychol*. 2017;31(1):59-84. doi:10.1080/13854046.2016.1190405
 28. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999;56(3):303-308. doi:10.1001/archneur.56.3.303
 29. Petersen RC, Doody R, Kurz A, et al. Current concepts in mild cognitive impairment. *Arch Neurol*. 2001;58(12):1985-1992. doi:10.1001/archneur.58.12.1985
 30. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256(3):183-194. doi:10.1111/j.1365-2796.2004.01388.x
 31. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256(3):240-246. doi:10.1111/j.1365-2796.2004.01380.x
 32. American Psychiatric Association, 2000. *Diagnostic and Statistical Manual of Mental Disorders IV Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association.
 33. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944. doi:10.1212/wnl.34.7.939
 34. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005
 35. Saxton J, Morrow L, Eschman A, Archer G, Luther J, Zuccolotto A. Computer assessment of mild cognitive impairment. *Postgrad Med*. 2009;121(2):177-185. doi:10.3810/pgm.2009.03.1990
 36. Tornatore JB, Hill E, Laboff JA, McGann ME. Self-administered screening for mild cognitive impairment: initial validation of a computerized test battery. *J Neuropsychiatry Clin Neurosci*. 2005;17(1):98-105. doi:10.1176/jnp.17.1.98
 37. Ahmed S, de Jager C, Wilcock G. A comparison of screening tools for the assessment of mild cognitive impairment: preliminary findings. *Neurocase*. 2012;18(4):336-351. doi:10.1080/13554794.2011.608365
 38. Scanlon L, O'Shea E, O'Caioimh R, Timmons S. Usability and validity of a battery of computerised cognitive screening tests for detecting cognitive impairment. *Gerontology*. 2016;62(2):247-252. doi:10.1159/000433432
 39. Gualtieri CT, Johnson LG. Neurocognitive testing supports a broader concept of mild cognitive impairment. *Am J Alzheimers Dis Other Dement*. 2005;20(6):359-366. doi:10.1177/153331750502000607
 40. Gualtieri CT, Johnson LG. Reliability and validity of a computerized neurocognitive test battery. *CNS Vital Signs*. *Arch Clin Neuropsychol*. 2006;21(7):623-643. doi:10.1016/j.acn.2006.05.007
 41. Dougherty JH Jr, Cannon RL, Nicholas CR, et al. The computerized self test (CST): an interactive, internet accessible cognitive screening test for dementia. *J Alzheimers Dis*. 2010;20(1):185-195. doi:10.3233/JAD-2010-1354
 42. de Jager CA, Schrijnemaekers AC, Honey TE, Budge MM. Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerized test battery, the Hopkins Verbal Learning Test and the MMSE. *Age Ageing*. 2009;38(4):455-460. doi:10.1093/ageing/afp068
 43. Hammers D, Spurgeon E, Ryan K, et al. Reliability of repeated cognitive assessment of dementia using a brief computerized battery. *Am J Alzheimers Dis Other Dement*. 2011;26(4):326-333. doi:10.1177/1533317511411907
 44. Lim YY, Jaeger J, Harrington K, et al. Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian Imaging, Biomarkers, and Lifestyle-rate of change substudy (AIBL-ROCS). *Arch Clin Neuropsychol*. 2013;28(4):320-330. doi:10.1093/arclin/act021
 45. Hammers D, Spurgeon E, Ryan K, et al. Validity of a brief computerized cognitive screening test in dementia. *J Geriatr Psychiatry Neurol*. 2012;25(2):89-99. doi:10.1177/0891988712447894
 46. Maruff P, Lim YY, Darby D, et al. Clinical utility of the CogState brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer's disease. *BMC Psychol*. 2013;1(1):30. doi:10.1186/2050-7283-1-30
 47. Mielke MM, Machulda MM, Hagen CE, et al. Performance of the CogState computerized battery in the Mayo Clinic Study on Aging. *Alzheimers Dement*. 2015;11(11):1367-1376. doi:10.1016/j.jalz.2015.01.008
 48. Fredrickson J, Maruff P, Woodward M, et al. Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology*. 2010;34(2):65-75. doi:10.1159/000264823
 49. Jacova C, McGrenere J, Lee HS, et al. C-TOC (Cognitive Testing on Computer): investigating the usability and validity of a novel self-administered cognitive assessment tool in aging and early dementia. *Alzheimer Dis Assoc Disord*. 2015;29(3):213-221. doi:10.1097/WAD.0000000000000055
 50. Scharre DW, Chang SI, Nagaraja HN, Vretos NE, Bornstein RA. Digitally translated Self-Administered Gerocognitive Examination (eSAGE): relationship with its validated paper version, neuropsychological evaluations, and clinical assessments. *Alzheimers Res Ther*. 2017;9(1):44. doi:10.1186/s13195-017-0269-3
 51. Kluger BM, Saunders LV, Hou W, et al. A brief computerized self-screen for dementia. *J Clin Exp Neuropsychol*. 2009;31(2):234-244. doi:10.1080/13803390802317559
 52. Cohen J, 1988. *Statistical power analysis for the behavioral sciences*, second edition. Hillsdale, NJ: Lawrence Erlbaum Associates.
 53. Millett G, Naglie G, Upshur R, Jaakkimainen L, Charles J, Tierney MC. Computerized cognitive testing in primary care: a qualitative study. *Alzheimer Dis Assoc Disord*. 2018;32(2):114-119. doi:10.1097/WAD.0000000000000219
 54. Sabbag MN, Boada M, Borson S, et al. Early detection of mild cognitive impairment (MCI) in primary care. *J Prev Alzheimers Dis*. 2020;7(3):165-170. doi:10.14283/jpad.2020.21
 55. Tierney MC, Naglie G, Upshur R, Moineddin R, Charles J, Jaakkimainen RL. Feasibility and validity of the self-administered computerized assessment of mild cognitive impairment with older primary care patients. *Alzheimer Dis Assoc Disord*. 2014;28(4):311-319. doi:10.1097/WAD.0000000000000036
 56. Adler K, Apple S, Friedlander A, et al. Computerized cognitive performance assessments in the Brooklyn Cognitive Impairments in Health Disparities Pilot Study. *Alzheimers Dement*. 2019;15(11):1420-1426. doi:10.1016/j.jalz.2019.07.004
 57. Babulal GM, Quiroz YT, Albeni BC, et al. Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimers Dement*. 2019;15(2):292-312. doi:10.1016/j.jalz.2018.09.009
 58. Mielke MM, Machulda MM, Hagen CE, et al. Influence of amyloid and APOE on cognitive performance in a late middle-aged cohort. *Alzheimers Dement*. 2016;12(3):281-291. doi:10.1016/j.jalz.2015.09.010
 59. Papp KV, Rentz DM, Maruff P, et al. The Computerized Cognitive Composite (C3) in an Alzheimer's Disease Secondary Prevention Trial. *J Prev Alzheimers Dis*. 2021;8(1):59-67. doi:10.14283/jpad.2020.38.