



Published in final edited form as:

*Alzheimers Dement.* 2015 November ; 11(11): 1367–1376. doi:10.1016/j.jalz.2015.01.008.

## Performance of the CogState computerized battery in the Mayo Clinic Study on Aging

Michelle M. Mielke<sup>a,b,\*</sup>, Mary M. Machulda<sup>c</sup>, Clinton E. Hagen<sup>d</sup>, Kelly K. Edwards<sup>d</sup>, Rosebud O. Roberts<sup>a,b</sup>, V. Shane Pankratz<sup>e</sup>, David S. Knopman<sup>b</sup>, Clifford R. Jack<sup>f</sup>, and Ronald C. Petersen<sup>b</sup>

<sup>a</sup>Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, 55905, USA

<sup>b</sup>Department of Neurology, Mayo Clinic, Rochester, MN, 55905, USA

<sup>c</sup>Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, 55905, USA

<sup>d</sup>Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, 55905, USA

<sup>e</sup>Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, 87131, USA

<sup>f</sup>Department of Radiology, Mayo Clinic, Rochester, MN, 55905, USA

### Abstract

**Background**—The feasibility and validity of brief computerized cognitive batteries at the population-level are unknown.

**Methods**—Non-demented participants ( $n = 1660$ , age 50–97) in the Mayo Clinic Study on Aging completed the computerized CogState battery and standard neuropsychological battery. The correlation between tests was examined and comparisons between CogState performance on the personal computer (PC) and iPad ( $n = 331$ ), and in the Clinic vs. at home ( $n = 194$ ), were assessed.

**Results**—We obtained valid data on >97% of participants on each test. Correlations between the CogState and neuropsychological tests ranged from  $-0.462$  to  $0.531$ . While absolute differences

\*Corresponding author: Michelle M. Mielke, Ph.D., Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, 200 First Street S.W., Rochester, MN 55905 USA. Tel.: +507-284-5545; Fax: +507-284-1516., mielke.michelle@mayo.edu.

#### Conflict of Interest Disclosures

Dr. Mielke served as a consultant to Eli Lilly and AbbVie, and receives research support from the NIH/NIA, the Alzheimer's Drug Discovery Foundation, and the Michael J. Fox Foundation. Dr. Machulda, Mr. Hagen, Ms. Edwards, Dr. Roberts, and Dr. Pankratz report no disclosures. Dr. Knopman serves as Deputy Editor for *Neurology*<sup>®</sup>; served on a Data Safety Monitoring Board for Lilly Pharmaceuticals; served as a consultant to TauRx Pharmaceuticals; was an investigator in clinical trials sponsored by Baxter, Elan Pharmaceuticals, and Forest Pharmaceuticals in the past 2 years; and receives research support from the NIH. Dr. Jack provides consulting services for Janssen Research & Development, LLC. He receives research funding from the National Institutes of Health, and the Alexander Family Alzheimer's Disease Research Professorship of the Mayo Foundation. Dr. Petersen serves on scientific advisory boards for Pfizer, Inc., Janssen Alzheimer Immunotherapy, Roche, Inc., Merck, Inc., and Genentech, Inc.; receives royalties from the publication of *Mild Cognitive Impairment* (Oxford University Press, 2003); and receives research support from the NIH/NIA.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

between the PC and iPad were small and participants preferred the iPad, performance on the PC was faster. Participants performed faster on Detection, One Card Learning, and One Back at home compared to the Clinic.

**Conclusions**—The computerized CogState battery, especially the iPad, was feasible, acceptable, and valid in the population.

### Keywords

Computerized cognitive battery; Epidemiology; Neuropsychology; Cognitively normal; Mild cognitive impairment; Population-based cohort study

## 1. Introduction

Alzheimer's disease (AD)-pathophysiology begins several years before the emergence of clinical symptoms [1,2]. It is therefore important to identify cognitive changes in the asymptomatic or early mild cognitive impairment (MCI) stages of AD. Standard pencil and paper neuropsychological testing is a critical element for the clinical diagnosis of dementia and determination of dementia type. However, traditional neuropsychological tests are labor-intensive, time-consuming, and associated with practice effects. A more efficient instrument is needed to serially assess cognition at the population-level. This is especially critical if secondary preventive trials prove effective.

Computerized testing may be better suited as a cognitive screening tool in large epidemiologic studies and for longitudinal monitoring by primary care providers. A computerized battery may have advantage over standard neuropsychological tests or other cognitive screening measures (e.g., Mini-Mental State Examination) by being more sensitive and efficient, removing ceiling and floor effects, providing real-time data entry and precise recording of accuracy and speed of response, minimizing practice effects, and suitability for off-site or long-distance use [3–6]. However, studies of computerized cognitive batteries have mainly been conducted in the clinical research setting using selected volunteers. This is important because there is still concern that computerized cognitive testing may not be feasible for elderly individuals and those with low education or little computer experience in the general population.

We incorporated the CogState computerized cognitive battery into the Mayo Clinic Study on Aging (MCSA). We included the CogState battery because it is brief, requires minimal administrative oversight, has a web-based platform, is easy to understand for individuals with little computer experience (e.g., [4,7,8]), and has good test-retest reliability (e.g., [4,9,10]). CogState is also being used as an endpoint in the Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4) Trial and the Dominantly Inherited Alzheimer's Network Trials (DIAN-TU) prevention trials.

In the present analysis, we had several aims. First, we characterized the feasibility of the CogState computerized battery in individuals aged 50–97 and determined factors that were associated with the inability to take the test. Second, we characterized CogState performance by diagnosis (normal cognition vs. MCI), APOE E4 genotype, age, and sex. Third, while previous studies provided correlations between the CogState tests and neuropsychological

tests [8,11,12], correlations between the CogState tests and a more extensive neuropsychological battery are limited. Therefore, we provided correlations between CogState tests and the standard neuropsychological test components administered in the MCSA. Lastly, among a subset of individuals, we also administered CogState on an iPad and at home. We compared performance on these different platforms and described our experiences.

## 2. Methods

### 2.1. Participants

The MCSA is a study of cognitive aging among Olmsted County, MN, residents that began in October 2004, and initially enrolled individuals aged 70 to 89 years. Follow-up visits were conducted every 15 months. The details of the study design and sampling procedures have been previously published [13]. Given the importance of understanding risk factors for the development and progression of AD pathophysiology in middle age, we expanded the study in 2012 to also enroll a population-based sample of individuals aged 50–69 using the same stratified random sampling methodology as in the original cohort. The present analysis includes 1,660 non-demented individuals aged 50–97 who completed both the CogState computerized battery and the standard neuropsychological battery at the same study visit.

### 2.2. Standard protocol approvals, registrations, and patient consents

The study protocols were approved by the Mayo Clinic and Olmsted Medical Center Institutional Review Boards. All participants provided written informed consent.

### 2.3. Participant assessment

MCSA study visits included a neurologic evaluation by a physician, an interview by a study coordinator, and neuropsychological testing by a psychometrist [13]. The physician examination included a medical history review, complete neurological examination, and administration of the Short Test of Mental Status [14] and the Unified Parkinson's Disease Rating Scale [15]. The study coordinator interview included demographic information and medical history, administration of the Beck Depression Inventory and the Beck Anxiety Inventory, and questions about memory to both the participant and an informant using the Clinical Dementia Rating scale [16]. The neuropsychological battery included nine tests covering four domains: 1) **memory** (Auditory Verbal Learning Test Delayed Recall Trial [17], Wechsler Memory Scale-Revised Logical Memory II & Visual Reproduction II) [18]; 2) **language** (Boston Naming Test [19] and Category Fluency [20]); 3) **executive function** (Trail Making Test B [21] and WAIS-R Digit Symbol subtest [22]); and 4) **visuospatial skills** (WAIS-R Picture Completion and Block Design subtests) [22]. Blood was collected and APOE genotype was determined. Medical comorbidities were assessed using information from the medical record to compute the Charlson Index [23].

### 2.4. Diagnostic determination

For each participant, performance in a cognitive domain was compared with the age-adjusted scores of cognitively normal individuals previously obtained using Mayo's Older American Normative Studies [24]. This approach relies on prior normative work and

extensive experience with the measurement of cognitive abilities in an independent sample of subjects from the same population. Subjects with scores of 1.0 SD below the age-specific mean in the general population were considered for possible cognitive impairment, taking into account education, prior occupation, visual or hearing deficits, and other information. A final decision to diagnose MCI was based on a consensus agreement between the interviewing nurse, examining physician, and neuropsychologist, after a review of all participant information [13,25]. Individuals who performed in the normal range and did not meet criteria for MCI or dementia were deemed cognitively normal. Performance on the computerized cognitive battery was not used to obtain a diagnosis. Individuals with dementia were not administered the CogState computerized battery.

## 2.5. CogState computerized battery administration in the Clinic

Administration of the CogState battery during the study visit was conducted on a personal computer (PC). The battery included four card tasks and the Groton Maze Learning Test (GMLT) [4,26–28]. Given the previous literature showing an initial practice/learning effect between the first and second administration [10], we administered a short practice battery, followed by a two-minute rest period, then the complete battery. The study coordinator was available to help the participants understand the tasks during the practice session. During the test battery, the coordinator provided minimal supervision or assistance. The tests were administered in this order:

*Detection (DET) task* – a simple reaction time paradigm that measures psychomotor speed. Reaction time (in milliseconds) was the primary outcome measure.

*Identification (IDN) task* – a choice reaction time paradigm that measures visual attention. Reaction time was the primary outcome measure.

*One Card Learning (OCL) task* – a continuous visual recognition learning task that assesses learning and attention. Reaction time and accuracy were the primary outcome measures.

*One Back (ONB) task* – a task that assesses working memory. Reaction time was the primary outcome measure.

*GMLT* – a hidden pathway maze learning test that measures spatial working memory, learning efficiency, and error monitoring [27,28]. The primary outcome measures were the average correct moves per second (speed/efficiency) across the five trials and the total number of errors.

The CogState battery provides a large number of equivalent alternative forms. This is achieved by having a large stimulus set from which exemplars are randomly chosen at run time, resulting in a different set of exemplars that are used each time an individual takes the test. The paradigm remains constant but the items are randomly chosen. Furthermore, the correct response (either yes or no) is randomly chosen for each trial of the task, and the inter-stimulus interval has a random interval that varies for each trial of the task. For the GMLT there are 20 possible hidden pathways matched for number of tiles and turns. These are presented in a random, non-recurring order which allows for 20 equivalent alternative forms.

## 2.6. Administration of CogState on the iPad and at home

We compared performance on the iPad and PC CogState versions in a subset of 341 individuals. The batteries were administered consecutively at the same study visit, with a 2–3 minute break between. The order of administration was alternated.

We also piloted a home-based battery using the four card tasks. We did not include the GMLT because it is more difficult and may require more administrative oversight. We compared performance on the home-based battery and clinic PC battery, taken within 6 months, among 194 individuals. We sent an email to participants explaining the study and providing a link with embedded de-identified study numbers so that they did not have to remember or find their ID number to take the test.

## 2.7. Response input for the in-clinic, iPad, and at home CogState versions

The response input for the CogState tasks varied by mode of administration. For the in-clinic and at-home versions, the keyboard was used for the four card tasks. A mouse was used for the in-clinic GMLT. On the iPad, the input was by finger touch for the four card tasks and by stylus for the GMLT.

## 2.8. Statistical methods

The ability to adhere to the requirements of each task, which could result from lack of understanding, cognitive impairment, or inattention to directions, was determined by integrity checks. Per the test developers and previous publications, data on each specific test were determined invalid if the accuracy of: DET was <90%, IDN was <80%, OCL was <50%, ONB was 70%. The GMLT data were considered invalid if the test was not completed within 10 minutes [4,28]. Because the raw data from the CogState tests are skewed, the data outputted are automatically transformed using logarithmic base 10 transformation for reaction time data and arcsine transformation for accuracy data to normalize the variables (see [4,10,26,28] for additional description). Chi-square tests for categorical variables and Kruskal-Wallis or equal variance two sample t-tests for continuous variables were used to examine differences between groups or modalities as appropriate. Pearson correlations were used to assess the relationship between the CogState tests and individual neuropsychological test z-scores. The z-scores for this correlation were not age-adjusted since the CogState data is not age-adjusted.

## 3. Results

Of the 1,660 participants, 1,574 were cognitively normal and 86 had MCI (see Table 1 for participant characteristics). As expected, individuals with MCI were older, had lower education, more depressive and anxiety symptoms, more medical comorbidities, and performed worse on standard neuropsychological tests (Table 1).

### 3.1. Factors associated with meeting integrity criteria for each CogState test

Of the participants taking each CogState test, 2.7% did not meet integrity criteria (i.e., had invalid data) on the DET, 1.7% on IDN, 2.1% on OCL, 3.0% on ONB, and 1.3% on the GMLT. The characteristics of the individuals who did and did not meet integrity criteria are

shown in Table 2. Across all tests, individuals who did not meet integrity were older, had worse cognitive performance on the standard neuropsychological battery, had more medical comorbidities, and more frequently had a diagnosis of MCI compared to those that did. However, the majority of individuals with MCI (>75% for all but the GMLT) were able to complete the CogState tests with valid data. Individuals with lower education were less likely to meet the integrity criteria on more complex tests of visual learning, working memory and problem-solving (OCL, ONB, and GMLT), but not on tests of simple and choice reaction times (DET, IDN). Greater depressive and anxiety symptoms were associated with the inability to meet integrity criteria on the ONB. APOE E4 genotype was not associated with the inability to meet integrity criteria on any test. Some individuals with cataracts or other vision problems were not able to complete the GMLT.

### 3.2. CogState performance by MCI diagnosis, age, sex, and APOE E4 genotype

After excluding individuals who did not meet integrity criteria on each test, individuals with MCI performed worse on all CogState tests ( $P < .001$ ; Table 3) and also took longer to complete the tests (34.4 (12.5) vs. 28.4 (9.6) minutes,  $P < .001$ ). Participants >65 years old (vs. 65 years) also performed worse, even after excluding those with MCI, but the mean differences were negligible. Women performed worse than men on the GMLT with more total errors (median [IQR]: 56 [42.0, 69.5] vs. 49 [38.0, 62.5],  $P < .001$ ) and slower speed (median [IQR]: 0.5 [0.4, 0.6] vs. 0.6 [0.4, 0.7],  $P = .006$ ). There were no differences in performance on any test by APOE E4 genotype.

### 3.3. Correlations between CogState tests and tests from the standard neuropsychological battery

Pearson correlations between each CogState test and global-, and domain-specific, and individual test z-scores are shown in Table 4. We initially examined the correlations separately by cognitive diagnosis and age (50–69 years vs. 70 and older). The correlations were similar across all subsets so we included everyone in the final analysis. Given the large sample size, all associations were  $P < .001$  with correlations ranging from  $-0.462$  to  $0.531$ . The only exception was the Boston Naming Test ( $\rho = -0.046$ ,  $P = .067$ ). However, many of the correlations were low.

### 3.4. PC vs. iPad comparison

Among a subset of 341 participants, we compared CogState performance on the PC versus iPad. There were no differences, by platform, in the percent of individuals with invalid data. Compared to the iPad, individuals performed faster on the PC (Table 5). They were also slightly more accurate on the OCL for the PC. However, the differences, while significant, were small. Our experience was that the majority of participants preferred the iPad over the PC and thought they did better on the iPad. Only ~2% of participants preferred the PC over the iPad.

### 3.5. PC vs. Home comparison

We also compared CogState performance on the four card tasks on the PC when taken in the Clinic versus at home on a PC in a subset of 194 individuals aged 50–69. We initially



administered the home battery to this younger age group because we thought they would feel more comfortable with the computer and capable of completing the battery without oversight. Compared to taking the test in the Clinic, participants were faster at home but accuracy did not differ (Table 6).

The feasibility of the home-based test depended on the internet provider and the participant's computer experience. For example, the link to the test embedded in the email wrapped around or did not work for some providers. This situation was particularly difficult for individuals with little computer knowledge who did not understand how to copy and paste the link into the address bar.

#### 4. Discussion

With the increasing recognition of the need to detect preclinical cognitive changes, it is critical to identify a more efficient means of assessing cognition at the population level. While standard neuropsychological tests are important for determining the clinical diagnosis of dementia, they are time-consuming, labor-intensive, and not suitable for repeated testing at the population-level [4,29]. Brief computerized cognitive batteries may be a better option for this purpose, particularly if they are easy to understand and designed for repeated administration (i.e., minimal practice effects and lack of ceiling effects). Several computerized tests have been developed including CogState, CANTAB, CNS Vital Signs, and NIH Toolbox (see [3], [5], and [6] for a more in-depth comparison of computerized tests). In the present study, we characterized the feasibility of the computerized CogState battery in the population-based MCSA, identified factors associated with test completion, compared performance to a standard neuropsychological battery, and compared performance on different platforms.

The feasibility of the computerized tests was demonstrated by the ability to obtain valid data for most individuals (97% for each test), including those with MCI. Older individuals or those with MCI were less likely to have valid data, but the overall percentage of invalid data for any given test was low. As the MCSA is population-based, we administered CogState to individuals who had never used a computer or cell phone, with low education (e.g., elementary education only), psychiatric disorders (axis I and axis II), multiple medical comorbidities, and over the age of 90, demonstrating that the test is feasible in the general population. While individuals in Olmsted County, MN, aged 50 and older, are primarily Caucasian, our results may not be generalizable to other ethnicities. However, our experience is consistent with the successful administration of the tests to indigenous Australians (e.g., [30,31]), and adults and children with schizophrenia, traumatic brain injury, acute stroke, and HIV [8,32,33].

Older individuals performed statistically worse than younger individuals (see Table 3), but this statistical difference was largely driven by the sample size since the absolute difference was very small. Regardless, this may suggest that some elderly individuals need more administrative oversight and help in understanding the test directions and/or navigating the computer. As the current middle-agers, with more computer experience, become elderly this difference may become less apparent. We also observed a higher frequency of invalid data

amongst those with lower education, possibly reflecting less exposure to computers, and individuals with medical comorbidities. In our experience, individuals with vision problems had difficulties with the lines in the GMLT maze, and individuals with arthritis or tremor had difficulty with the mouse. Other variables including sex, depression, anxiety, and APOE E4 genotype had little effect on test performance.

Consistent with previous studies, individuals with MCI performed worse on all CogState tests [11,26,34,35] and were more likely to have invalid data. However, most individuals with a consensus diagnosis of MCI were still able to complete many of the tests in the clinic. We have not examined whether it is feasible, and the data valid, for MCI participants to self-administer the tests without any oversight, but this presumably would be more difficult. We also did not test individuals with a dementia diagnosis, but previous studies have administered the tests to patients with mild forms of dementia [34]. The benefit of CogState is likely in detecting subtle cognitive changes in preclinical states, and may have little use in patients with frank dementia.

The four card tasks assess psychomotor speed (DET), visual attention (IDN), visual learning (OCL), and working memory (ONB) [8]. The GMLT assesses spatial problem solving [36]. The CogState tasks were chosen by the test developers because they assess aspects of cognitive function that are commonly affected in neurodegenerative and neuropsychiatric conditions. The tasks were further optimized for brief administration, minimization of practice effects, and the availability of outcome measures that could be used for parametric analyses [8,36]. As such, they differ from traditional psychometric measures used in clinical practice. The CogState tasks significantly correlated with virtually all the standard neuropsychological tests and domain scores though the strength of the correlation varied. There were moderate correlations between standard neuropsychological tests of attention/executive function (i.e., WAIS-R Digit Symbol; Trailmaking Test B) and the CogState IDN, ONB, and GMLT. There was also a moderate correlation between WAIS-R Block Design and the GMLT. However, the remaining correlations were in the weak to very weak range. This indicates that CogState tests do not map one-to-one to standard neuropsychological tests. The low correlations do not mean that the CogState tests are not valid. It is possible that the CogState tests provide greater measurement sensitivity for cognitive change among cognitively normal individuals, and these analyses are ongoing.

The mode of response input (e.g., keyboard, mouse, finger or stylus touch) can affect both the speed and accuracy of the results [37]. When we compared performance on the iPad vs. the PC, individuals were faster and more accurate using the keyboard on the PC compared to the finger touch on the iPad. For the GMLT, individuals were also faster with a mouse on the PC than on the iPad with a stylus touch. The observed difference across platforms is consistent with Wood and colleagues [37] who reported that a mouse yielded more accurate results than a touchscreen. Thus, the response input is important to consider when comparing test mediums and interpreting results. Ideally, if taken to the clinic, the response input would need to be standardized. Despite performing slightly faster and more accurately on the PC, participants preferred the iPad over the PC and also think they do better on the iPad. This observation was particularly true for the elderly, especially those with little PC experience and with arthritis or tremors. Therefore, despite some performance differences on



the iPad versus the PC, the iPad was more feasible for, and accepted by, older individuals in the population.

We initially thought that individuals would perform worse at home compared to in the clinic because at home they did not have anyone to help explain the instructions or answer any questions. In contrast, the participants performed faster at home. When asked about their testing experience, these participants often reported they would take the tests when they felt their freshest, typically mid-morning. This might explain the slightly faster performance at home and highlights the need to consider the time of day the test is taken in relation to performance. In the present study, participants all completed the computerized test in the clinic before they took the test at home. Therefore, the results may differ if individuals took the test at home first or are test naive. In contrast to some individuals doing better at home because they are maximizing their time of peak performance, others may do worse due to the potential for distractions in the environment. A critical step in validating computerized tests will be to better understand and determine how to control for environmental variability and testing supervision. Notably, a recent study described the acceptability and feasibility of monthly administration, over one year, of the four card tasks solely through an online at-home battery in 143 individuals with no previous exposure to the test [29]. While only 5% of individuals at baseline were not able to complete the test, the participants were recruited from the community and had to meet eligibility requirements. Another study of similar design with a larger group of individuals is ongoing ([www.brainhealthregistry.com](http://www.brainhealthregistry.com)). These studies highlight the possibility of at-home testing, but our experience is that this will only work for some individuals at the population-level with sufficient computer knowledge and access.

With the increasing use of technology in the world, the use of computerized cognitive testing is likely to change the way we longitudinally monitor cognition in the future for both research and clinical purposes. Standard pencil and paper neuropsychological testing is a critical element for the clinical diagnosis of dementia and determination of dementia type and will likely remain the gold standard for diagnostic purposes. However, with the increasing need to identify cognitive changes as early as possible in the population, when individuals are still considered cognitively normal, a less time- and energy-intensive instrument is needed. As such, computerized testing may be better suited as a cognitive screening tool in large epidemiologic studies, for monitoring individuals in a registry for AD drug trials, and for longitudinal monitoring by primary care providers. There are many different types of computerized cognitive testing but it is unlikely that one test will serve all needed purposes (e.g., detecting cognitive change, detecting or diagnosing MCI, sensitivity to a specific neuroimaging or CSF biomarker). Our results suggest that computerized testing is feasible but additional longitudinal comparisons of the feasibility, acceptability, and performance across environments is needed. Further, direct longitudinal comparisons of cognitive performance on computerized batteries with known neuroimaging and CSF biomarkers of dementia are needed to fully elucidate and interpret the test results in relation to the disease progression of AD and other dementias.

## Acknowledgments

This study was supported by funding from the National Institutes of Health/National Institute on Aging P50 AG016574, U01 AG006786, R01 AG041851, and R01 AG011378; the Robert H. and Clarice Smith and Abigail van Buren Alzheimer's Disease Research Program, the Walter S. and Lucienne Driskill Foundation, and was made possible by the Rochester Epidemiology Project (R01 AG034676). The funding organizations did not have a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

## Abbreviations

<b>AD</b>	Alzheimer's disease
<b>APOE</b>	apolipoprotein E
<b>AVLT</b>	Auditory Verbal Learning Task
<b>AVLT DR</b>	Auditory Verbal Learning Task, Delayed Recall
<b>DET</b>	detection
<b>GMLT</b>	Groton Maze Learning Test
<b>HV</b>	hippocampal volume
<b>IDN</b>	identification
<b>IQR</b>	interquartile range
<b>MCI</b>	mild cognitive impairment
<b>MCSA</b>	Mayo Clinic Study of Aging
<b>OCL</b>	One Card Learning
<b>ONB</b>	One Back
<b>PC</b>	personal computer
<b>WMS-R LM II</b>	Wechsler Memory Scale-Revised Logical Memory II
<b>WMS-R VR II</b>	Wechsler Memory Scale-Revised Visual Reproduction II

## References

1. Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* 1991; 82:239–59. [PubMed: 1759558]
2. Shaw LM, Korecka M, Clark CM, Lee VM, Trojanowski JQ. Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. *Nat Rev Drug Discov.* 2007; 6:295–303. [PubMed: 17347655]
3. Snyder PJ, Jackson CE, Petersen RC, Khachaturian AS, Kaye J, Albert MS, et al. Assessment of cognition in mild cognitive impairment: a comparative study. *Alzheimer's Dement.* 2011; 7:338–55. [PubMed: 21575877]
4. Fredrickson J, Maruff P, Woodward M, Moore L, Fredrickson A, Sach J, et al. Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology.* 2010; 34:65–75. [PubMed: 20016215]
5. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. *Alzheimer's Dement.* 2008; 4:428–37. [PubMed: 19012868]
6. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. *Am J Alzheimers Dis Other Dement.* 2014;10.1177/1533317514522852

7. Dingwall KM, Lewis MS, Maruff P, Cairney S. Reliability of repeated cognitive testing in healthy Indigenous Australian adolescents. *Aust Psychol.* 2009; 44:224–34.
8. Maruff P, Thomas E, Cysique L, Brew B, Collie A, Snyder P, et al. Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Arch Clin Neuropsychol.* 2009; 24:165–78. [PubMed: 19395350]
9. Falletti MG, Maruff P, Burman P, Harris A. The effects of growth hormone (GH) deficiency and GH replacement on cognitive performance in adults: a meta-analysis of the current literature. *Psychoneuroendocrinology.* 2006; 31:681–91. [PubMed: 16621325]
10. Collie A, Maruff P, Darby DG, McStephen M. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *J Int Neuropsychol Soc.* 2003; 9:419–28. [PubMed: 12666766]
11. de Jager CA, Schrijnemaekers AC, Honey TE, Budge MM. Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. *Age Ageing.* 2009; 38:455–60. [PubMed: 19454402]
12. Tippett WJ, Lee JH, Mraz R, Zakzanis KK, Snyder PJ, Black SE, et al. Convergent validity and sex differences in healthy elderly adults for performance on 3D virtual reality navigation learning and 2D hidden maze tasks. *Cyberpsychol Behav.* 2009; 12:169–74. [PubMed: 19361297]
13. Roberts RO, Geda YE, Knopman DS, Cha RH, Pankratz VS, Boeve BF, et al. The Mayo Clinic Study of Aging: design and sampling, participation, baseline measures and sample characteristics. *Neuroepidemiology.* 2008; 30:58–69. [PubMed: 18259084]
14. Kokmen E, Smith GE, Petersen RC, Tangalos E, Ivnik RC. The short test of mental status. Correlations with standardized psychometric testing. *Arch Neurol.* 1991; 48:725–8. [PubMed: 1859300]
15. Fahn, SER. Committee MotUD. Unified Parkinson's Disease Rating Scale. Florham Park: MacMillan Healthcare Information; 1987.
16. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology.* 1993; 43:2412–4. [PubMed: 8232972]
17. Rey, A. *L'examen Clinique en Psychologie.* Paris: Presses Universitaires de France; 1964.
18. Wechsler, D. *Manual for the Wechsler Memory Scale-Revised.* San Antonio, TX: The Psychological Corporation; 1987.
19. Kaplan, E.; Goodglass, H.; Weintraub, S. *The Boston Naming Test.* Philadelphia: Lea & Febiger; 1983.
20. Strauss, E.; Sherman, EMS.; Spreen, O. *A Compendium of Neuropsychological Tests.* New York: Oxford University Press; 2006.
21. Reitan R. Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills.* 1958; 8:271–6.
22. Wechsler, D. *Wechsler Adult Intelligence Scale-Revised [Manual].* San Antonio, TX: Psychological Corporation; 1981.
23. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol.* 1992; 45:613–9. [PubMed: 1607900]
24. Ivnik RJ, Malec JF, Smith GE, Tangalos EG, Petersen RC, Kokmen E, et al. Mayo's Older Americans Normative Studies: WAIS-R norms for ages 56 to 97. *Clin Neuropsychol.* 1992; 6:1–30.
25. Petersen RC, Roberts RO, Knopman DS, Geda YE, Cha RH, Pankratz VS, et al. Prevalence of mild cognitive impairment is higher in men: the Mayo Clinic Study of Aging. *Neurology.* 2010; 75:889–97. [PubMed: 20820000]
26. Lim YY, Ellis KA, Harrington K, Ames D, Martins RN, Masters CL, et al. Use of the CogState Brief Battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. *J Clin Exp Neuropsychol.* 2012; 34:345–58. [PubMed: 22248010]
27. Snyder PJ, Bednar MM, Cromer JR, Maruff P. Reversal of scopolamine-induced deficits with a single dose of donepezil, an acetylcholinesterase inhibitor. *Alzheimer's Dement.* 2005; 1:126–35. [PubMed: 19595845]

28. Pietrzak RH, Maruff P, Mayes LC, Roman SA, Sosa JA, Snyder PJ. An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Arch Clin Neuropsychol*. 2008; 23:433–45. [PubMed: 18448309]
29. Darby DG, Fredrickson J, Pietrzak RH, Maruff P, Woodward M, Brodtmann A. Reliability and usability of an Internet-based computerized cognitive testing battery in community-dwelling older people. *Comput Human Behav*. 2014; 30:199–205.
30. Cairney S, Clough A, Jaragba M, Maruff P. Cognitive impairment in Aboriginal people with heavy episodic patterns of alcohol use. *Addiction*. 2007; 102:909–15. [PubMed: 17523985]
31. Pearce MS, Mann KD, Singh G, Sayers SM. Birth weight and cognitive function in early adulthood: the Australian Aboriginal birth cohort study. *J Dev Orig Health Dis*. 2014; 5:240–7. [PubMed: 24901664]
32. Cumming TB, Brodtmann A, Darby D, Bernhardt J. Cutting a long story short: reaction times in acute stroke are associated with longer term cognitive outcomes. *J Neurol Sci*. 2012; 322:102–6. [PubMed: 22819353]
33. Boivin MJ, Busman RA, Parikh SM, Bangirana P, Page CF, Opoka RO, et al. A pilot study of the neuropsychological benefits of computerized cognitive rehabilitation in Ugandan children with HIV. *Neuropsychology*. 2010; 24:667–73. [PubMed: 20804255]
34. Hammers D, Spurgeon E, Ryan K, Persad C, Barbas N, Heidebrink J, et al. Validity of a brief computerized cognitive screening test in dementia. *J Geriatr Psychiatry Neurol*. 2012; 25:89–99. [PubMed: 22689701]
35. Papp KV, Snyder PJ, Maruff P, Bartkowiak J, Pietrzak RH. Detecting subtle changes in visuospatial executive function and learning in the amnesic variant of mild cognitive impairment. *PLoS ONE*. 2011; 6:e21688. [PubMed: 21779333]
36. Pietrzak RH, Cohen H, Snyder PJ. Spatial learning efficiency and error monitoring in normal aging: an investigation using a novel hidden maze learning test. *Arch Clin Neuropsychol*. 2007; 22:235–45. [PubMed: 17306502]
37. Wood E, Willoughby T, Rushing A, Bechtel L, Gilbert J. Use of computer input devices by older adults. *J Applied Gerontol*. 2005; 24:419–38.

### Research in context

- 1. Systematic review:** We reviewed the literature in PubMed examining the feasibility of the CogState computerized and in relationship to standard neuropsychological testing. Studies of computerized cognitive batteries have mainly been conducted in the clinical research setting using selected volunteers. There are few, if any, population-based studies and none that compare performance on a PC versus iPad or at home versus in the Clinic.
- 2. Interpretation:** The results of our study support the feasibility and acceptability of the CogState tests in a population-based study of individuals aged 50–97. While the frequency of invalid data was low overall (<3% on any test), individuals who were older, had MCI, or lower education were more likely to have invalid data. There were some moderate correlations between the CogState tests and standard neuropsychological tests, but most correlations were weak. This indicates that the CogState tests do not map one-to-one to standard neuropsychological tests. Participants performed slightly faster on the PC than the iPad, but the iPad was preferred, especially among the elderly, and may be more amenable to testing in the general population. Participants performed slightly better at home than in the Clinic.
- 3. Future Directions:** Continued follow-up of our cohort will help to determine the longitudinal change in the CogState tests in relation to risk of MCI, dementia, and changes in neuroimaging measures of AD pathology. We will also further assess at-home testing in a larger, population-based cohort.

Table 1

Participant characteristics by cognitive status

Characteristic	Cognitively normal (CN)		Mild cognitive impairment (MCI)		P-value
	N	Median (IQR)/N(%)	N	Median (IQR)/N(%)	
Age	1574	66 (60, 74)	86	68 (64, 77)	.008
Women	1574	768 (48.7%)	86	41 (47.7%)	.84
Education, years	1574	15 (13, 17)	86	13 (12, 15)	<.0001
APOE E4 allele	1427	396 (27.8%)	84	28 (33.3%)	.184
Body Mass Index, kg/m <sup>2</sup>	1574	28.4 (25.4, 32.4)	86	30.2 (26.9, 34.1)	.011
Beck Depression Inventory	1574	3 (1.0, 6.0)	86	6 (3.0, 11.0)	<.0001
Beck Anxiety Inventory	1574	1 (0.0, 4.0)	86	4 (1.0, 9.0)	<.0001
Charlson Index	1574	2 (1.0, 4.0)	86	3 (2.0, 5.0)	<.0001
<i>Standard Neuropsychological Battery</i>					
MMSE*	1574	29 (28.0, 29.0)	86	26 (24.0, 27.0)	<.0001
Global Z-score	1574	0.5 (-0.1, 1.0)	86	-1.2 (-1.8, -0.6)	<.0001
Memory Z-score	1574	0.4 (-0.2, 1.0)	86	-1.1 (-1.9, -0.5)	<.0001
Language Z-score	1574	0.3 (-0.3, 0.8)	86	-0.9 (-1.6, -0.2)	<.0001
Attention Z-score	1574	0.4 (-0.1, 0.9)	86	-0.8 (-1.5, 0.0)	<.0001
Visual spatial Z-score	1574	0.3 (-0.3, 0.9)	86	-0.7 (-1.5, -0.0)	<.0001

Abbreviation: IQR, Interquartile range.

\* MMSE scores were modified from the Short Test of Mental Status.



Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Met criteria (N=1630)	Identification (IDN)			One Card Learning (OCL)			One Back (ONB)			Groton Maze Learning Test (GMLT)		
	Median (IQR)/N(%)	Did not meet criteria (N=28)	p-value	Met criteria (N=1622)	Did not meet criteria (N=35)	p-value	Met criteria (N=1605)	Did not meet criteria (N=49)	p-value	Met criteria (N=1500)	Did not meet criteria (N=19)	p-value
66 (60, 73)	75.5 (66, 82.5)	0.0002	66 (60, 74)	69 (65, 76)	0.008	66 (60, 73)	78 (69, 86)	<0.0001	65 (60, 70)	74 (64, 84)	0.0007	
795 (48.8%)	13 (46.4%)	0.806	790 (48.7%)	19 (45.8%)	0.514	780 (48.6%)	26 (53.1)	0.538	734 (48.9%)	9 (47.4%)	0.892	
15 (13, 17)	14 (12, 15.5)	0.109	15 (13, 17)	13 (12, 16)	0.002	15 (13, 17)	13 (12, 16)	0.009	15 (13, 16)	13 (12, 16)	0.044	
418/1479 (38.7%)	6/28 (21.4%)	0.426	412/1473 (28.0%)	11/32 (34.4%)	0.425	405/1455 (27.8%)	16/47 (34.0%)	0.351	372/1350 (27.6%)	5/18 (27.8%)	0.983	
28.4 (25.4, 32.5)	29.9 (27.1, 33.3)	0.117	28.4 (25.4, 32.4)	28.9 (26.4, 34.0)	0.270	28.4 (25.5, 32.4)	29.2 (25.2, 33.3)	0.334	28.5 (25.5, 32.6)	27.5 (23.0, 32.8)	0.324	
3 (1, 6)	3.5 (2, 8.5)	0.144	3 (1, 6)	4 (1, 8)	0.276	3 (1, 6)	7 (4, 9)	<0.0001	3 (1, 6)	3 (1, 9)	0.440	
1 (0, 4)	1 (0, 6)	0.489	1 (0, 4)	1 (0, 6)	0.480	1 (0, 4)	3 (1, 8)	0.011	1 (0, 4)	1 (0, 3)	0.729	
2 (1, 4)	3.5 (2, 6)	0.002	2 (1, 4)	3 (1, 5)	0.026	2 (1, 4)	4.5 (3, 6.5)	<0.0001	2 (1, 3)	4 (3, 6)	0.0002	
81 (5.0%)	5 (17.9%)	0.002	82 (5.1%)	4 (11.4%)	0.093	74 (4.6%)	10 (20.4%)	<0.0001	71 (4.7%)	6 (31.6%)	<0.0001	
29 (28, 29)	28 (26, 29)	0.003	29 (28, 29)	28 (27, 29)	0.002	29 (28, 29)	28 (26, 28.5)	<0.0001	29 (28, 29)	28 (25, 29)	0.001	
0.4 (-0.2, 0.9)	-0.7 (-1.5, -0.1)	<0.0001	0.4 (-0.2, 0.9)	-0.2 (-0.8, 0.4)	0.000	0.4 (-0.2, 0.9)	-0.4 (-1.3, 0.1)	<0.0001	0.5 (-0.1, 1.0)	-0.8 (-1.5, 0.0)	<0.0001	
0.4 (-0.3, 1.0)	-0.4 (-0.9, 0.4)	0.004	0.4 (-0.3, 1.0)	0 (-0.5, 0.5)	0.040	0.4 (-0.3, 1.0)	-0.2 (-0.9, 0.4)	0.0001	0.4 (-0.3, 1.0)	-0.3 (-1.1, 0.6)	0.018	
0.3 (-0.3, 0.8)	-0.4 (-1.2, 0.0)	0.0001	0.3 (-0.3, 0.8)	-0.2 (-0.9, 0.5)	0.005	0.3 (-0.3, 0.8)	-0.5 (-1.0, 0.3)	<0.0001	0.3 (-0.3, 0.8)	-0.4 (-1.3, -0.1)	0.0004	
0.3 (-0.2, 0.8)	-0.4 (-1.4, 0.2)	<0.0001	0.3 (-0.2, 0.8)	-0.5 (-1.3, 0.4)	<0.0001	0.4 (-0.2, 0.9)	-0.6 (-1.4, -0.0)	<0.0001	0.4 (-0.1, 0.9)	-1.0 (-1.5, -0.1)	<0.0001	
0.3 (-0.3, 0.9)	-0.4 (-1.2, 0.3)	0.0001	0.3 (-0.3, 0.9)	0 (-0.8, 0.6)	0.011	0.3 (-0.3, 0.9)	-0.3 (-0.9, 0.2)	<0.0001	0.3 (-0.3, 0.9)	-0.5 (-0.9, 0.4)	0.002	

Alzheimers Dement. Author manuscript; available in PMC 2016 November 01.

Table 3

CogState test performance by diagnosis and age

	CN			MCI			65 years			>65 years		
	N	Mean (SD)	P-value*	N	Mean (SD)	P-value*	N	Mean (SD)	P-value*	N	Mean (SD)	P-value*
DET - time	1534	2.6 (0.1)	<.0001	79	2.7 (0.1)	<.0001	753	2.6 (0.1)	<.0001	864	2.6 (0.1)	<.0001
IDN - time	1549	2.7 (0.1)	<.0001	81	2.8 (0.1)	<.0001	758	2.7 (0.1)	<.0001	877	2.8 (0.1)	<.0001
OCL - time	1540	1.0 (0.1)	<.0001	82	0.9 (0.1)	<.0001	756	1.0 (0.1)	<.0001	871	1.0 (0.1)	<.0001
OCL - accuracy	1540	3.0 (0.1)	<.0001	82	3.1 (0.1)	<.0001	756	3.0 (0.1)	<.0001	871	3.1 (0.1)	<.0001
ONB - time	1531	2.9 (0.1)	<.0001	74	3.0 (0.1)	<.0001	758	2.9 (0.1)	<.0001	852	3.0 (0.1)	<.0001
GMLT - errors	1429	53.9 (19.4)	<.0001	71	69.1 (18.9)	<.0001	754	49.6 (17.1)	<.0001	750	59.8 (20.7)	<.0001
GMLT - speed	1429	0.6 (0.1)	<.0001	71	0.4 (0.1)	<.0001	754	0.6 (0.1)	<.0001	750	0.5 (0.1)	<.0001

Abbreviations: CN, Cognitively Normal; SD, Standard Deviation, MCI, Mild Cognitive Impairment; DET, Detection; IDN, Identification; OCL, One Card Learning; ONB, One Back; GMLT, Groton Maze Learning Test.

\* All P-values were derived from two-sample t-tests.

Pearson correlations between the CogState tests and standard neuropsychological tests administered at the same study visit

**Table 4**

	DET time	IDN time	ONB time	OCL time	OCL accuracy	GMLT errors	GMLT speed
Global Z-score	-0.354	-0.383	-0.404	-0.282	0.298	-0.406	0.478
Memory Z-score	-0.188	-0.204	-0.219	-0.151	0.280	-0.294	0.230
WMS-R Logical Memory Delayed Recall	-0.109	-0.113	-0.134	-0.090	0.160	-0.191	0.126
AVLT Delayed Recall	-0.126	-0.158	-0.158	-0.134	0.263	-0.185	0.140
Visual reproduction Delayed Recall	-0.217	-0.220	-0.229	-0.136	0.248	-0.336	0.293
Executive Function Z-score	-0.389	-0.462	-0.510	-0.410	0.249	-0.334	0.531
Digit Symbol	-0.326	-0.418	-0.443	-0.366	0.200	-0.239	0.455
TMT, Part B	-0.358	-0.386	-0.447	-0.345	0.239	-0.357	0.472
Language Z-score	-0.238	-0.263	-0.240	-0.158	0.185	-0.228	0.314
Boston Naming Test	-0.188	-0.152	-0.127	-0.046*	0.090	-0.206	0.195
Category Fluency	-0.201	-0.271	-0.257	-0.201	0.204	-0.167	0.309
Visual spatial Z-score	-0.286	-0.261	-0.294	-0.171	0.184	-0.374	0.381
Picture Completion	-0.195	-0.172	-0.181	-0.091	0.090	-0.223	0.208
Block Design	-0.282	-0.261	-0.305	-0.189	0.212	-0.388	0.412

Abbreviations: DET, Detection; IDN, Identification; OCL, One Card Learning; ONB, One Back; GMLT, Groton Maze Learning Test, WMS-R, Wechsler Memory Scale-Revised; AVLT, Auditory Verbal Learning Test; TMT, Trail Making Test.

\*  $P$ -values for all Pearson correlations were  $P < .001$ , with the exception of the Boston Naming Test and OCL-time ( $P = .067$ ).

**Table 5**

Comparisons of PC and iPad CogState results administered in the Clinic

	N	Clinic PC mean (SD)	Clinic iPad mean (SD)	Mean Difference PC vs. iPad mean (SD)	Mean Difference 95% CI	t	P-value
DET - time	331	2.593 (0.107)	2.598 (0.107)	-0.005 (0.094)	-0.015, 0.005	-1.02	.308
IDN - time	341	2.716 (0.062)	2.741 (0.072)	-0.025 (0.053)	-0.031, -0.020	-8.74	<.001
OCL - time	339	3.008 (0.076)	3.017 (0.081)	-0.009 (0.048)	-0.014, -0.004	-3.48	<.001
OCL - accuracy	339	1.057 (0.090)	1.028 (0.089)	0.029 (0.097)	0.018, 0.039	-5.47	<.001
ONB - time	326	2.895 (0.078)	2.900 (0.080)	-0.005 (0.060)	-0.012, 0.001	-1.53	.128
GMLT - errors	338	45.423 (15.320)	46.716 (16.358)	-1.293 (14.971)	-2.895, 0.309	-1.59	.113
GMLT - speed	338	0.620 (0.144)	0.774 (0.180)	-0.154 (0.133)	-0.168, -0.140	-21.27	<.001

Abbreviations: PC, personal computer; SD, standard deviation; DET, Detection; IDN, Identification; OCL, One Card Learning; ONB, One Back; GMLT, Groton Maze Learning Test.

\* Analyses were conducted using paired t-tests.

**Table 6**  
Comparisons of the CogState results on the PC administered in the Clinic versus at home

	N	Clinic PC mean (SD)	Home PC mean (SD)	Mean Difference PC vs. Home* mean (SD)	Mean Difference 95% CI	t	P-value
DET - time	188	2.600 (0.116)	2.574 (0.093)	0.025 (0.104)	0.010, 0.040	3.51	.001
IDN - time	194	2.721 (0.060)	2.718 (0.063)	0.003 (0.057)	-0.005, 0.011	0.77	.441
OCL - time	192	3.009 (0.084)	2.976 (0.075)	0.034 (0.053)	0.026, 0.041	8.91	<.001
OCL - accuracy	192	1.052 (0.101)	1.056 (0.087)	-0.004 (0.095)	-0.018, 0.009	-0.61	.543
ONB - time	185	2.891 (0.087)	2.864 (0.088)	0.027 (0.058)	0.018, 0.035	6.24	<.001

Abbreviations: PC, personal computer; SD, standard deviation; DET, Detection; IDN, Identification; OCL, One Card Learning; ONB, One Back; GMLT, Groton Maze Learning Test.

\* Analyses were conducted using paired t-tests.