#### Original Paper

# The Rapid Online Cognitive Assessment for the Detection of Neurocognitive Disorder: Open-Label Study

Calvin Howard<sup>1,2,3</sup>, MD; Amy Johnson<sup>2</sup>, BS; Joseph Peedicail<sup>2</sup>, MD; Marcus C Ng<sup>2,4</sup>, MD

#### **Corresponding Author:**

Calvin Howard, MD
Department of Neurology, Brigham and Women's Hospital
Harvard Medical School
60 Fenwood Road 1st Floor
Boston, MA, 02115
United States
Email: <a href="mailto:choward12@bwh.harvard.edu">choward12@bwh.harvard.edu</a>

### **Abstract**

**Background:** The rising prevalence of dementia necessitates a scalable solution to cognitive screening. Paper-based cognitive screening examinations are well-validated but minimally scalable. If a digital cognitive screening examination could replicate paper-based screening, it may improve scalability while potentially maintaining the performance of these well-validated paper-based tests. Here, we evaluate the Rapid Online Cognitive Assessment (RoCA), a remote and self-administered digital cognitive screening examination.

**Objective:** The objective of this study was to validate the ability of RoCA to reliably evaluate patient input, identify patients with cognitive impairment relative to the established tests, and evaluate its potential as a screening tool.

**Methods:** RoCA uses a convolutional neural network to evaluate a patient's ability to perform common cognitive screening tasks: wireframe diagram copying and clock drawing tests. To evaluate RoCA, we compared its evaluations with those of established paper-based tests. This open-label study consists of 46 patients (age range 33-82 years) who were enrolled from neurology clinics. Patients completed the RoCA screening examination and either Addenbrooke's Cognitive Examination-3 (ACE-3, n=35) or Montreal Cognitive Assessment (MoCA, n=11). We evaluated 3 primary metrics of RoCA's performance: (1) ability to correctly evaluate patient inputs, (2) ability to identify patients with cognitive impairment compared to ACE-3 and MoCA, and (3) performance as a screening tool.

**Results:** RoCA classifies patients similarly to gold standard paper-based tests, with a receiver operating characteristic area under the curve of 0.81 (95% CI 0.67-0.91; P<.001). RoCA achieved sensitivity of 0.94 (95% CI 0.80-1.0; P<.001). This was robust to multiple control analyses. Approximately 83% (16/19) of the patient respondents reported RoCA as highly intuitive, with 95% (18/19) perceiving it as adding value to their care.

**Conclusions:** RoCA may act as a simple and highly scalable digital cognitive screening examination. However, due to the limitations of this study, further work is required to evaluate the ability of RoCA to be generalizable across patient populations, assess its performance in an entirely remote manner, and analyze the effect of digital literacy.

J Med Internet Res 2025;27:e66735; doi: 10.2196/66735

**Keywords:** cognitive; neurology; dementia; geriatrics; artificial intelligence; assessment; online assessment; cognitive assessment; cognitive Assessment; cognitive Assessment; RoCA; cohort

<sup>&</sup>lt;sup>1</sup>Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

<sup>&</sup>lt;sup>2</sup>Section of Neurology, Department of Internal Medicine, University of Manitoba, Winnipeg, MB, Canada

<sup>&</sup>lt;sup>3</sup>Klinik für Neurologie mit Experimenteller Neurologie, Charité – Universitätsmedizin Berlin, Berlin, Germany

<sup>&</sup>lt;sup>4</sup>Graduate Program in Biomedical Engineering, University of Manitoba, Winnipeg, MB, Canada

# Introduction

Current projections estimate 150 million patients with dementia worldwide by 2050, with 57 million as of 2019 [1]. This causes considerable health care system strain, leaving a majority of patients undiagnosed [2-5]. However, for the patients who do receive a diagnosis, it often takes 3 years or longer from symptom onset [2,3,6-8]. The next step, that is, receiving an etiological diagnosis like Alzheimer disease, requires even more time [4,8].

Much research has previously focused on evaluating the contributors to these problems [9-13]. Frontline physicians report two key addressable issues: (1) the logistical difficulty of screening enough patients [9,10] and (2) variable comfort in diagnosing patients with dementia [6,14].

Digital cognitive assessments (DCAs) offer a promising solution to these issues [15-17]. They provide high scalability, which addresses logistical difficulties, and can render expert-level diagnoses, which address the issue of diagnostic comfort. However, most DCAs have limitations preventing them from completely addressing these two issues [18]. First and foremost is that these novel tests lack validation [15]. DCAs often use completely new testing methods, diverging from the well-validated methods that made paper-based tests so valuable [15-17,19,20]. The extensively validated nature of these tests helps define the patient populations and use-cases for these tests, while the creation of novel unrelated tests requires repetition of this process. It may be possible to piggyback new-age cognitive tests upon established tests, helping them achieve the generalizability that required decades of work and iteration to accomplish with established tests. Second, poor design choices often reduce accessibility [16,17,21-25]. Common examples are burying test access deep within websites, requiring users to make accounts, or having patients use unfamiliar hardware. Lastly, most DCAs are not truly scalable [16,17,21-25]. Despite being digital, they often require specific tablets, computers, on-site testing, or even expert test evaluators—these choices bottleneck patient access [19,26,27].

The Rapid Online Cognitive Assessment (RoCA) is a DCA that aims to address these limitations. To stay consistent with well-validated methods, it reproduces the screening results of gold-standard paper-based tests: Addenbrooke's Cognitive Examination-3 (ACE-3) and the Montreal Cognitive Assessment (MoCA) [28-31]. Patient input directly guided its design to ensure accessibility, resulting in a short touchscreen-based drawing battery with automated convolutional neural network—based scoring [32,33]. For scalability, RoCA is entirely automated, remote, functions on all devices, and utilizes cloud computing to enhance geographic access. Finally, RoCA aims to specifically act as a screening examination. For this reason, we prioritized its sensitivity.

We first ensured that RoCA's underlying machine learning works well by assessing the accuracy of its neural network. Then, we compared RoCA's similarity to gold standard paper-based tests and evaluated its accessibility

computationally and with patient input. Lastly, we provide the data-driven thresholds that maximize its sensitivity, optimizing its function as a screening examination [34,35].

#### Methods

#### Ethical Considerations

This study was conducted in accordance with ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments. Approval was received from the Research Ethics Board of the Bannatyne Campus, University of Manitoba (#HS25666). Patients were recruited from neurology and neuropsychology clinics, and all signed institutional review board–approved consent forms. Substitute decision makers were included in the consent process of patients with cognitive impairment. Patients did not receive compensation for involvement in the study. This work was conducted between October 01, 2022 and December 01, 2024.

#### Study Participants

Our study cohort enrolled patients from neurology clinics across the Health Sciences Center, University of Manitoba (n=46). Patients with and without cognitive impairment were recruited. The inclusion criterion was English fluency. The exclusion criteria were acute psychiatric disorder contributing to cognitive state, disability restricting ability to utilize screens, disability restricting ability to receive visual and auditory instructions, developmental delay, acute medical condition contributing to cognitive state, and specifically delirium. Patients indicating interest in clinical research were contacted by study team members via phone.

Interested patients were screened for inclusion and exclusion criteria and enrolled. At the first clinic visit, patients were again screened for inclusion or exclusion criteria by a physician. Patients or their caregivers provided written consent at the first clinic visit. Patients were recruited until sample size for statistical power was achieved.

#### Study Design

Patients were tested in a quiet environment by a physician trained in cognitive examination. RoCA was completed on a touchscreen tablet. RoCA automatically administered instructions to the patient and was completed automatically without interference or prompting from the examiner. The responses were automatically scored and summated without staff involvement. During RoCA, the patient was observed, but there was no interference from study staff in evaluation or scoring. ACE-3 and MoCA were administered and scored according to standard guidelines by one of the 3 trained experts [30,31]. Caregivers were allowed to join but could not participate in the examination.

# Cognitive Status Classification

A trained clinician administered a label of cognitive impairment based on the established cutoffs for each test: 26/30 on MoCA and 83/100 on ACE-3 [30,31,36].

#### Sample Size Calculation

Patients were enrolled based on sample size requirements defined by the Hanley and McNeil formula [37]. This formula is based on the area under the curve (AUC) of the receiver operating characteristic and describes how sample size requirements vary with AUC. Under relatively good performance, an AUC of 0.70 and 80% statistical power is achieved with 16 positive and 16 negative cases. Under optimal performance with an AUC of 0.90, 80% statistical power is achieved with 2 positive and 2 negative cases. We conservatively aimed to recruit 16 participants with cognitive impairment and 16 without cognitive impairment.

## RoCA Cognitive Screening Examination

RoCA is a self-administering cognitive screening examination, which is compatible with smartphones, tablets, and personal computers. It relies upon devices having an internet connection to ensure all patients can access it, regardless of hardware specifications or specific device.

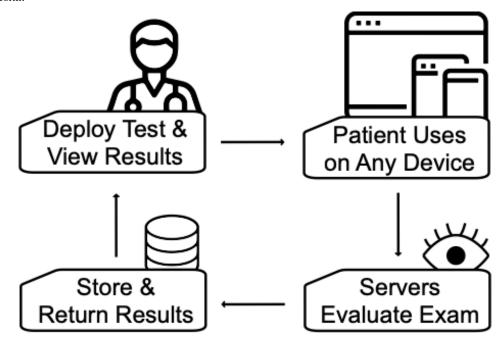
RoCA consists of 3 questions. Similar to previously described batteries, patients are asked to copy a line diagram of a cube, copy overlapping infinities, and perform a clock

drawing [32]. For each question, the patient has an unrestricted amount of time to answer. Each question's instructions are provided via a closed-captioned audio. Instructions may be repeated up to 3 times, but no further assistance is provided. Questions are answered via a touchscreen, although a keyboard and mouse may be used. A correct cube drawing is worth 2 points, overlapping infinities 1 point, and a clock 5 points. Incorrect drawings are worth 0 points. The total possible score is 8 points.

### RoCA Deployment

The RoCA evaluation starts with the clinician (Figure 1). The clinician uses an administrative platform, from which they send tests and view results, to send an encrypted access link to a patient. Patients open this link to begin RoCA. Links are specific to each given patient and are inactivated after use. Upon completing the test, the results are encrypted and sent to a scoring server on a private subnet. The scoring server sends encrypted scores to an encrypted database on a private subnet. The administrative platform receives scores from this database, allowing the clinician to view the results upon completion. The system is Health Information Privacy Protection Act—compliant.

Figure 1. The Rapid Online Cognitive Assessment (RoCA) deployment system. Clinicians begin the deployment system. They can use the administrative platform to deploy an access link to patients. Patients subsequently receive the access link and are then able to take the interactive test on any device via the internet. Once the patient completes the examination, their answers are passed to the servers, which then evaluate the patient's RoCA. The results are then stored in a database and are available for the patient's clinician to see. Clinicians can view a patient's results from the administrative platform.



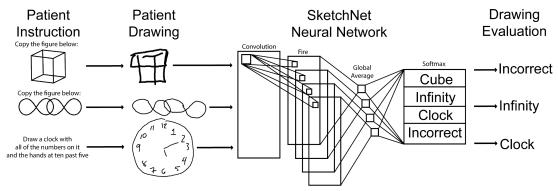
#### Patient Drawing Classification

Patient drawings were evaluated using SketchNet, a convolutional neural network built specifically to evaluate RoCA inputs (Figure 2) [33,38]. Briefly, the SketchNet is a convolutional neural network using a SqueezeNet architecture [39]. This particular architecture is composed of convolutional layers, fire modules, and finishes with a global average and softmax. This allows high degrees of accuracy while

maintaining the speed of classifications and a small overall size of the model. It was trained using transfer learning, pretrained on ImageNet, and subsequently trained on thousands of RoCA-specific drawings to evaluate cognitive test drawings with 97% accuracy [38,39].

SketchNet-based classifications were compared to ground truth: assessment of drawings by a clinician trained in cognitive examinations. The drawings were scored according to established scoring guidelines [30].

**Figure 2.** The Rapid Online Cognitive Assessment (RoCA) drawing evaluation system. Patients receive audiovisual instructions asking them to draw 3 different images. The first 2 tasks are image-copying tasks of a wire cube and overlapping infinities. The last question is to draw a clock face at 10 minutes past 5 PM. The patient drawings are then retrieved and preprocessed. After preprocessing, the images are then passed to SketchNet, a convolutional neural network, to classify the images. The output of SketchNet is the classification of the input image. The SketchNet is then applied to each image in RoCA.



# Evaluation of Drawing Classification

The SketchNet is inherently a multiclass classifier, which yields complex classification evaluations. To simplify this, we condensed classifications into correct versus incorrect. Confusion matrices were constructed from SketchNet classification outputs. Using the confusion matrix, we derived all classification metrics: accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. This was repeated for each drawing.

# Comparison of Drawing Classification Against a Statistical Baseline

We next investigated if RoCA was classifying images better than that expected by chance. To do this, we developed a random classifier, which represents classification performance at chance level. It is calculated by deriving a confusion matrix (true positives, true negatives, false positives, false negatives) under chance circumstances. The confusion matrix is thus the probability of selecting a given class multiplied by the probability of an image being a given class (equation S1 in Multimedia Appendix 1). All classification metrics then follow from this confusion matrix. These provide the chance-level baseline.

To compare the classification of each drawing, we bootstrapped SketchNet classifications (n=1000) and counted the number of times the bootstrap fell below chance level. This is the P value. To compare the overall performance, we averaged performance across RoCA and the chance-level performance and then compared them with an independent 2-sided t test.

# Evaluation of Diagnostic Classification

To evaluate the reliability of RoCA patient classifications, we compared RoCA classification to ground-truth classification. Ground-truth classification is the classification adjudicated by ACE-3 or MoCA in accordance with the established guidelines [30,31,36]. A receiver operating characteristic was constructed, and its AUC was calculated to measure diagnostic performance. The Youden Index was calculated to derive the optimal classification threshold [40]. Patient

classifications were based upon this threshold, and these were then used to construct confusion matrices for RoCA.

## Statistical Evaluation of RoCA Classification

To derive a statistical baseline for RoCA, we again employed a random classifier equivalent to RoCA. The confusion matrix was again generated using probabilities of selecting each class, as described above (equation S1 in Multimedia Appendix 1). The random-chance baseline for the AUC was chosen to be 0.50, in accordance with the literature [37]. To compare RoCA to these chance-level baselines, we again used the bootstrapping technique described above (n=10,000).

#### Classification Confidence

To evaluate the confidence of RoCA classifications, we estimated this directly with the confidence intervals derived from bootstrapping (n=10,000) [41,42]. At all possible RoCA diagnostic thresholds, we derived the confusion matrices and classification metrics for each bootstrap. This allows the observation of RoCA confidence across all possible thresholds. To make sure that an appropriate threshold is chosen to optimize RoCA for a screening examination, we focused on evaluating the sensitivity and negative predictive value across all thresholds.

#### Covariates Influencing the RoCA Score

We aimed to identify the clinical covariates that might be influencing RoCA scores. To do this, we collected several covariates: age, ethnicity, sex, educational status, employment status, and which paper-based examination they received. These covariates were then related to RoCA scores. This was done using a multivariate regression of all variables upon the RoCA score.

We also assessed if any individual covariate compounded the effect of impaired cognition. This was done with a series of additional regressions. In these, the covariate, cognitive status, and their interactions were regressed upon the RoCA score. This was done for each covariate.

# RoCA Usability and Patient Perception

A follow-up survey was sent to patients within 6 months of completing RoCA. Patients responded to a battery of questions by using both dichotomous (yes or no) questions and Likert scale questions. Likert scales were adjudicated such that 1 corresponded to very low, 2 was low, 3 was moderate, 4 was high, and 5 was very high.

#### Statistical Analysis

All analyses were performed in Python. The 2-sided *t* testing was performed with SciPy [43]. Scikit-learn was used for receiver operating characteristic curve construction [44]. Regression analyses were performed with statsmodels and an ordinary least squares regression [45]. Confidence

intervals were calculated using bootstrapping with replacement (n=1000) [41].

#### Results

#### Patient Characteristics

In this study, 143 patients were assessed for eligibility; 79 patients declined enrollment, 7 did not meet the inclusion criteria, 57 were enrolled, and 11 did not make their appointment. Finally, 46 patients completed the study. There were 16 patients with cognitive impairment and 30 patients without cognitive impairment. Patient demographics are shown in Table 1.

**Table 1.** Patient demographics (N=46)

Table 1. Patient demographics (N=46).	
Parameter	Values
Age (years), mean (SD)	49.1 (15.0)
Sex, n (%)	
Female	24 (52)
Male	22 (48)
Educational status, n (%)	
Less than secondary	2 (4)
Postsecondary	8 (17)
Secondary	36 (78)
Employment status, n (%)	
Employed	29 (63)
Unemployed	17 (37)
Ethnicity, n (%)	
African	1 (2)
Caucasian	31 (67)
European	1 (2)
Filipino	3 (7)
Indian	4 (9)
Indigenous	6 (13)
Cognitive status, n (%)	
Impaired	17 (37)
Intact	29 (63)
Diagnosis, n (%)	
Neurologically healthy	16 (35)
Mild cognitive impairment	9 (20)
Probable Alzheimer disease	7 (15)
Epilepsy	14 (30)
Cognitive examination, n (%)	
ACE-3 <sup>a</sup>	35 (76)
MoCA <sup>b</sup>	11 (24)

<sup>&</sup>lt;sup>a</sup>ACE-3: Addenbrooke's Cognitive Examination-3.

# RoCA's Evaluation of Patient Drawings

We first evaluated how RoCA evaluated patient drawings (Figure 3A-C). RoCA classified 97% (44/46) of the cubes

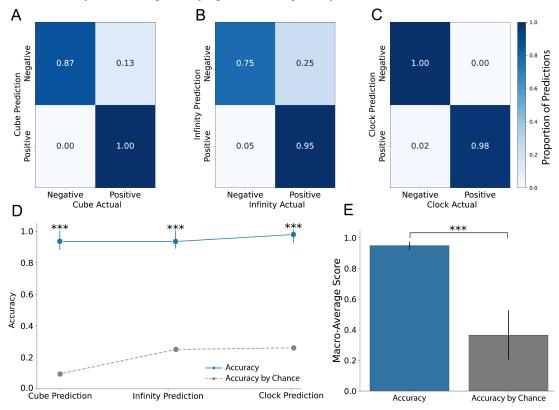
correctly, 91% (42/46) of the infinities correctly, and 98% (45/46) of the clocks correctly. We next calculated the accuracy of RoCA for each drawing individually (Figure 3D). We compared the accuracy of each drawing to its

<sup>&</sup>lt;sup>b</sup>MoCA: Montreal Cognitive Assessment.

statistical baseline by bootstrapping, resampling the accuracy, and counting the number of times it fell below the random classifier. The accuracy for the cube was 93% (95% CI 0.85-1.0; *P*<.001), for the overlapping infinities was 94% (95% CI 0.87-1.0; *P*<.001), and for the clock was 98% (95% CI 0.90-1.0; *P*<.001). Finally, we derived the overall

accuracy of RoCA across all drawings (Figure 3E). RoCA had 95% (SD 3%) accuracy across all drawings, which was higher than that expected by chance (*P*<.001). Additional classification metrics are available in Table S1 of Multimedia Appendix 1.

Figure 3. Accurate classification of patient drawings by the Rapid Online Cognitive Assessment (RoCA). (A) Confusion matrix for cube classification. (B) Confusion matrix for infinity classification. (C) Confusion matrix for clock classification. (D) Accuracy of RoCA compared to that expected by chance for all drawings. Bootstrapped confidence intervals were used to statistically compare the observed accuracy of RoCA to that expected by chance. (E) The overall accuracy of RoCA is significantly higher than that expected by chance.

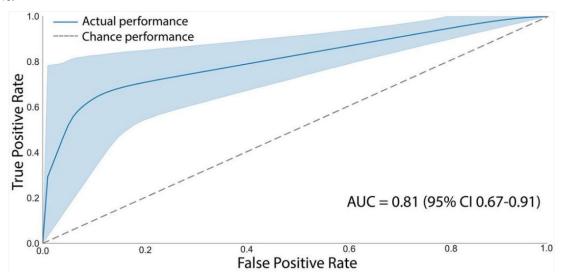


# RoCA Achieves Similar Diagnostic Fidelity to Gold-Standard Tests

Next, we compared the overall performance of RoCA to that of the gold-standard tests (Figure 4). Each patient's RoCA score was related to their diagnostic classification by using a receiver operating characteristic curve compared to ACE-3 and MoCA classifications. The AUC was 0.81, which was

significantly higher than that expected by chance (95% CI 0.67-0.91; P<.001). RoCA was found to outperform patient classification expected by chance ( $\chi^2_1$ =1.7; P<.001). This was robust, regardless of whether ACE-3 (AUC=0.79, 95% CI 0.67-0.86; P<.001) or MoCA (AUC=1.0, 95% CI 1.0-1.0; P<.001) was used. Subsequently, we found the optimal threshold for RoCA was 7/8, according to Youden Index [46].

**Figure 4.** Diagnostic performance of the Rapid Online Cognitive Assessment (RoCA) compared to that of the gold-standard paper-based test. The receiver operating characteristic of RoCA achieved a value of 0.81 (95% CI 0.67-0.91; *P*<.001). The receiver operating characteristic curve was calculated by bootstrapping RoCA's overall classification to derive confidence intervals and mean performance. The AUC presented is the mean AUC across all bootstraps. The shaded region represents 95% CI. The point-estimate AUC or the AUC without any bootstrap is 0.85. AUC: area under the curve.



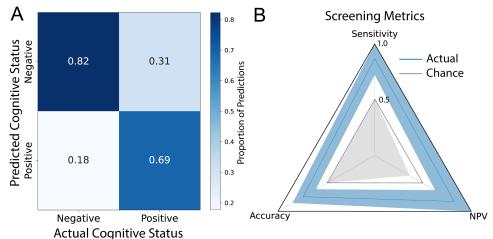
# Screening Performance of RoCA

We next evaluated RoCA's ability to act as a screening examination. To do this, we evaluated RoCA's accuracy, sensitivity, and negative predictive value. We began by developing a random classifier equivalent of RoCA, which we used to derive the RoCA's statistical baseline for comparison (Figure S1 in Multimedia Appendix 1). All screening metrics were expected to be low by chance, with an expected accuracy of 50%, sensitivity of 50%, and negative predictive value of 63%.

Following this, we evaluated the actual RoCA's screening performance. We began by calculating the confusion matrix

of the actual RoCA by using the optimal threshold of 7/8 (Figure 5A). We then calculated the screening metrics for RoCA (Figure 5B). At the optimal threshold, RoCA has an accuracy of 0.76 (95% CI 0.63-0.89; *P*<.001), which was better than that expected by chance. It also achieved both superior and statistically significant sensitivity of 0.94 (95% CI 0.81-1.0; *P*<.001) and a statistically significant negative predictive value of 0.95 (95% CI 0.84-1.0; *P*<.001). Specificity and positive predictive value were also calculated for completeness, although they are not directly related to the screening ability (Table S2 in Multimedia Appendix 1).

Figure 5. Screening performance of the Rapid Online Cognitive Assessment (RoCA). (A) Confusion matrix for the classification of RoCA at the threshold, normalized by predictions. The true positive rate was 62% (16/26), true negative rate was 95% (19/20), false negative rate was 5% (1/20), and false positive rate was 38% (10/26). (B) Screening metrics of RoCA presented in a radar plot. The solid blue vertices represent the measured screening metric, with shaded blue edges marking the 95% CIs. The gray interior represents the expected performance by random chance. RoCA achieved an excellent sensitivity of 0.94 (95% CI 0.63-0.89; P<.001) and an excellent negative predictive value of 0.95% CI 0.63-0.89; P<.001). NPV: negative predictive value.

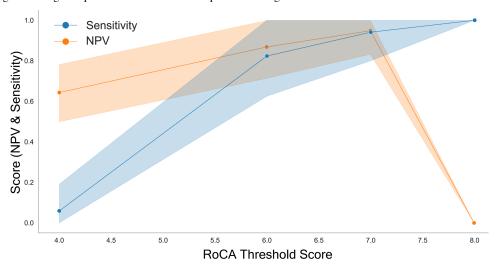


#### Diagnostic Confidence of RoCA

Next, we ensured that the chosen RoCA threshold is the optimal threshold for screening. Youden J, calculated using ROC, balances sensitivity and specificity, and therefore will not necessarily result in the optimal screening threshold. To search for the optimal screening threshold, we calculated the confidence of sensitivity and negative predictive value across all potential scores (Figure 6). We found that the threshold of 7/8, identified by AUC, was the optimal screening

threshold. It simultaneously maximized the sensitivity (0.94) and negative predictive value (0.95) while also optimizing the confidence interval for sensitivity (95% CI 0.81-1.0) and negative predictive value (95% CI 0.84-1.0). Specificity and positive predictive value were also calculated for completeness and are available, although these are not optimized in screening examinations (Figure S2 in Multimedia Appendix 1).

**Figure 6.** At the optimal threshold of 7/8, RoCA optimizes both sensitivity and NPV. At 7/8, the values of sensitivity and NPV are maximized with minimization of their uncertainty. At this point, sensitivity is also highly confident (sensitivity=0.94, 95% CI 0.80-1.0). NPV is also maximized (0.80, 95% CI 0.83-0.95). Shaded regions represent 95% CIs derived from bootstrapping (n=10,000). Points represent the estimated sensitivity and NPV without bootstrapping. NPV: negative predictive value. RoCA: Rapid Online Cognitive Assessment.



# Influence of Patient Demographics on RoCA Score

We next evaluated whether any patient factors may be influencing RoCA performance. We first performed one multivariate regression to evaluate the relationship of all the demographic variables with the RoCA score (Figure S2 in Multimedia Appendix 1). However, only cognitive status had a significant association with the RoCA score ( $\beta$ =1.07; P<.001). No other patient factor was related. We also performed a series of multivariate regressions for each covariate, assessing if patient factors might compound the effects of impaired cognition (Figure S3 in Multimedia Appendix 1). Again, we found that only cognitive status was significantly associated with RoCA scores. Lastly, we evaluated the time investment required to complete RoCA. We found that RoCA takes roughly 2 minutes and 30 seconds (mean 148, SD 34 s).

# Patient Perspective of RoCA

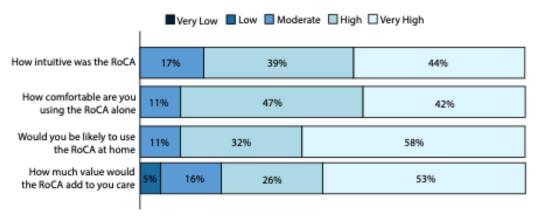
Finally, we evaluated accessibility as reported by patients. We did this using post-RoCA survey responses, wherein patients answered questions specifically regarding the accessibility of RoCA (Figure 7); 19 patients (mean age 60.2, SD 18 years)

of the 46 patients responded. The cognitive status of the respondents was unknown, given the anonymous nature of the survey.

Likert scales were used to derive an evaluation of the overall patient evaluation of RoCA. Approximately 83% (16/19) of the respondents reported RoCA as highly or very highly intuitive, 90% (17/19) of the respondents reported high or very high comfort in using RoCA, 90% (17/19) of the respondents reported they would be highly or very highly confident using the test alone, and 79% (15/19) of the respondents reported that RoCA would add a high or very high degree of value to their medical care.

We also asked patients a series of yes or no questions regarding other aspects of RoCA (Table S3 in Multimedia Appendix 1). Among them, we found 100% of the patients would want to take the test prior to appointments to discuss results, and 100% of the patients would trust RoCA's results. Most surprisingly, we found 55% (10/19) of the patients made appreciable lifestyle changes after using RoCA. These changes specifically included starting cognitive exercises, beginning a dementia-friendly diet, starting physical exercise, or doing financial planning.

Figure 7. Patients' accessibility of RoCA. This is a bar plot of the survey responses to Likert scale questions from 19 patients post RoCA. RoCA: Rapid Online Cognitive Assessment.



#### Discussion

# Interpretation of Results

We found that RoCA performed well in this single-center open-label trial. RoCA can evaluate the drawings used in cognitive examinations and subsequently use them to sensitively screen for cognitive impairment. Importantly, RoCA does this in-line with established paper-based tests. Further, we found that RoCA is an accessible screening system, as we did not find significant effects of patient factors such as age or demographics. This combined with RoCA's cloud-based platform allows it to act as a sensitive, accessible, and scalable digital cognitive screening examination.

# Role of RoCA in Screening

The first limitation of DCAs is their lack of validity compared to paper-based screening systems [15-17,19,20]. RoCA specifically aimed to classify patients similarly as paper-based screening systems, thereby acting as a digital surrogate for them [28-31,47-49]. We found that RoCA rules out (screens) cognitively healthy patients, which would have been similarly ruled out by standard paper-based tests.

The primary benefit of RoCA's sensitivity is triaging patients for further examination. It can prioritize at-risk patients while offloading the cognitively healthy for routine observation. This process eliminates the proportion of true negatives from the population, who go on for subsequent assessments, thereby increasing the positive predictive values and negative predictive values of any further evaluation [50]. RoCA may benefit individual clinical practices in triaging patients, or it could aid large-scale screening of patient populations.

#### RoCA in Context

Although the currently available list of digital cognitive examinations has been comprehensively reviewed elsewhere [15,16], we will briefly place RoCA in the context of other digital cognitive examinations. RoCA has several key distinguishing features. First, RoCA is entirely automated and self-administered, allowing it to be administered remotely. Other digital cognitive examinations such as MoCA Duo require an expert to physically administer the test

[19], similar to the paper-based testing. Second, RoCA is developed for accessibility. Most digital cognitive tests require patients to navigate complex websites [21], app stores [51], or use keyboards and mouses [52]. RoCA is developed such that patients simply receive an access link via text or email, click it, and complete the test on any device of their choosing. Finally, RoCA specifically emulates established paper-based tests. Other DCAs often develop new machine learning algorithms or cognitive evaluation maneuvers [51-53] and then relate these to specific diagnostic classifications. However, although these are potentially highly useful, the generalizability of these new tests is unclear and will require extensive additional studies to demonstrate generalizability and validity similar to paper-based tests.

#### Limitations

This study and RoCA are not without limitations. First, this study does not focus on individual etiologies causing dementia but focuses on identifying cognitive impairment as a whole. Thus, it is possible that etiologies presenting in different cognitive domains may result in variable RoCA performance. However, before specializing into the evaluation of different etiologies, it is critical to accurately screen cognitive impairment itself. This will help RoCA generalize across disorders causing cognitive impairment. Another limitation is the sample size. Although this study's sample size was defined using a power analysis, RoCA will require additional testing in larger cohorts across additional demographics, larger age ranges, specific neurodegenerative etiologies, and a variety of digital literacy levels. Additionally, further studies will need to be performed with blinding. Lastly, the survey results must be interpreted with caution. Due to the nature of surveys and the 6-month survey time frame, it is possible that the survey results are subject to biases such as the selection bias.

There are limitations to RoCA. First, RoCA is hardware-dependent. For patients without touchscreens, they may have difficulty in generating high-quality drawings, which may hinder performance. However, RoCA was specifically trained on a dataset mixing drawings generated from touchscreens, mouses, and stylets to specifically offset this risk. Beyond this, RoCA is internet connection–dependent. However, we have ensured that this is compatible with smartphones

to leverage their inherent internet connection. Although RoCA has been developed to be remotely administered and automatically scored, this study only evaluates the ability of a patient to perform RoCA alone, but still within the clinic. Thus, the robustness of RoCA under variable conditions in a fully remote setting must be evaluated in further works. It remains to be seen how stable RoCA scores are over time, and test-retest reliability must be assessed in future studies. Finally, RoCA is specifically a screening examination. It is not a test developed to have high specificity. Thus, to

consolidate a positive screening result for a diagnosis of cognitive impairment, a patient should receive a subsequent evaluation with high specificity. For these reasons, RoCA is best used in a 2-part system, wherein the first test prioritizes sensitivity, while the second prioritizes specificity. For example, RoCA could be followed by our other full-length diagnostic test, the Autonomous Cognitive Examination [54,55]. The utility of combining these 2 tests at scale will require further work.

#### **Acknowledgments**

The authors would like to acknowledge Rachel Elizabeth Bethune Howard for supporting this research.

#### **Conflicts of Interest**

CH is part of CogNet Inc, a company that provides cognitive testing to rural Canadians.

#### Multimedia Appendix 1

Supplementary data.

[DOCX File (Microsoft Word File), 780 KB-Multimedia Appendix 1]

#### References

- 1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health. Feb 2022;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
- 2. Boise L, Morgan DL, Kaye J, Camicioli R. Delays in the diagnosis of dementia: perspectives of family caregivers. Am J Alzheimers Dis (Columbia). Jan 1999;14(1):20-26. [doi: 10.1177/153331759901400101]
- 3. Draper B, Cations M, White F, et al. Time to diagnosis in young-onset dementia and its determinants: the INSPIRED study. Int J Geriatr Psychiatry. Nov 2016;31(11):1217-1224. [doi: 10.1002/gps.4430] [Medline: 26807846]
- 4. 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. Feb 3, 2017;7(2):e011146. [doi: 10.1136/bmjopen-2016-011146] [Medline: 28159845]
- 5. Goodman RA, Lochner KA, Thambisetty M, Wingo TS, Posner SF, Ling SM. Prevalence of dementia subtypes in United States Medicare fee-for-service beneficiaries, 2011-2013. Alzheimers Dement. Jan 2017;13(1):28-37. [doi: 10.1016/j.jalz.2016.04.002] [Medline: 27172148]
- 6. Morgan D, Kosteniuk J, O'Connell ME, et al. Barriers and facilitators to development and implementation of a rural primary health care intervention for dementia: a process evaluation. BMC Health Serv Res. Oct 17, 2019;19(1):709. [doi: 10.1186/s12913-019-4548-5] [Medline: 31623609]
- 7. Kvello-Alme M, Bråthen G, White LR, Sando SB. Time to diagnosis in young onset Alzheimer's disease: a population-based study from Central Norway. J Alzheimers Dis. 2021;82(3):965-974. [doi: 10.3233/JAD-210090] [Medline: 34120901]
- 8. Fiske A, Gatz M, Aadnøy B, Pedersen NL. Assessing age of dementia onset: validity of informant reports. Alzheimer Dis Assoc Disord. 2005;19(3):128-134. [doi: 10.1097/01.wad.0000174947.76968.74] [Medline: 16118529]
- 9. Dubois B, Padovani A, Scheltens P, Rossi A, Dell'Agnello G. Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. J Alzheimers Dis. 2016;49(3):617-631. [doi: 10.3233/JAD-150692] [Medline: 26484931]
- 10. Liss JL, Seleri Assunção S, Cummings J, et al. Practical recommendations for timely, accurate diagnosis of symptomatic Alzheimer's disease (MCI and dementia) in primary care: a review and synthesis. J Intern Med. Aug 2021;290(2):310-334. [doi: 10.1111/joim.13244] [Medline: 33458891]
- 11. Knopman D, Donohue JA, Gutterman EM. Patterns of care in the early stages of Alzheimer's disease: impediments to timely diagnosis. J Am Geriatr Soc. Mar 2000;48(3):300-304. [doi: <a href="https://doi.org/10.1111/j.1532-5415.2000.tb02650.x">10.1111/j.1532-5415.2000.tb02650.x</a>] [Medline: 10733057]
- 12. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. BMJ. Jun 16, 2015;350:h3029. [doi: 10.1136/bmj.h3029] [Medline: 26079686]
- 13. Iliffe S, Robinson L, Brayne C, et al. Primary care and dementia: 1. diagnosis, screening and disclosure. Int J Geriatr Psychiatry. Sep 2009;24(9):895-901. [doi: 10.1002/gps.2204] [Medline: 19226529]

- 14. 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] [Medline: 19568149]
- 15. Chan JYC, Yau STY, Kwok TCY, Tsoi KKF. Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: a systematic review. Ageing Res Rev. Dec 2021;72:101506. [doi: 10.1016/j.arr.2021.101506] [Medline: 34744026]
- 16. Cubillos C, Rienzo A. Digital cognitive assessment tests for older adults: systematic literature review. JMIR Ment Health. Dec 8, 2023;10:e47487. [doi: 10.2196/47487] [Medline: 38064247]
- 17. Magno M, Martins AI, Pais J, Silva AG, Rocha NP. Diagnostic accuracy of digital solutions for screening for cognitive impairment: a systematic review and meta-analysis. Appl Sci (Basel). 2024;14(6):2640. [doi: 10.3390/app14062640]
- 18. Tsoy E, Zygouris S, Possin KL. Current state of self-administered brief computerized cognitive assessments for detection of cognitive disorders in older adults: a systematic review. J Prev Alzheimers Dis. 2021;8(3):267-276. [doi: 10.14283/jpad.2021.11] [Medline: 34101783]
- 19. MoCA Cognition. URL: https://mocacognition.com/ [Accessed 2025-06-09]
- 20. Babrak LM, Menetski J, Rebhan M, et al. Traditional and digital biomarkers: two worlds apart? Digit Biomark. 2019;3(2):92-102. [doi: 10.1159/000502000] [Medline: 32095769]
- 21. Brain health assessment. Cogniciti. URL: <a href="https://cogniciti.com/Test-Your-Brain-Health/Brain-Health-Assessment">https://cogniciti.com/Test-Your-Brain-Health/Brain-Health-Assessment</a> [Accessed 2025-06-09]
- 22. Danousis M, Goumopoulos C. A machine-learning-based motor and cognitive assessment tool using in-game data from the GAME2AWE platform. Informatics. 2023;10(3):59. [doi: 10.3390/informatics10030059]
- 23. Javed AR, Fahad LG, Farhan AA, et al. Automated cognitive health assessment in smart homes using machine learning. Sustain Cities Soc. Feb 2021;65:102572. [doi: 10.1016/j.scs.2020.102572]
- 24. Rykov YG, Patterson MD, Gangwar BA, et al. Predicting cognitive scores from wearable-based digital physiological features using machine learning: data from a clinical trial in mild cognitive impairment. BMC Med. Jan 25, 2024;22(1):36. [doi: 10.1186/s12916-024-03252-y] [Medline: 38273340]
- 25. Almgren H, Camacho M, Hanganu A, et al. Machine learning-based prediction of longitudinal cognitive decline in early Parkinson's disease using multimodal features. Sci Rep. Aug 14, 2023;13(1):13193. [doi: 10.1038/s41598-023-37644-6] [Medline: 37580407]
- 26. Freedman M, Leach L, Carmela Tartaglia M, et al. The Toronto Cognitive Assessment (TorCA): normative data and validation to detect amnestic mild cognitive impairment. Alzheimers Res Ther. Jul 18, 2018;10(1):65. [doi: 10.1186/s13195-018-0382-y] [Medline: 30021658]
- 27. Ehrensperger MM, Taylor KI, Berres M, et al. BrainCheck a very brief tool to detect incipient cognitive decline: optimized case-finding combining patient- and informant-based data. Alzheimers Res Ther. 2014;6(9):69. [doi: 10.1186/s13195-014-0069-y] [Medline: 25422675]
- 28. Larner AJ, Mitchell AJ. A meta-analysis of the accuracy of the Addenbrooke's Cognitive Examination (ACE) and the Addenbrooke's Cognitive Examination-Revised (ACE-R) in the detection of dementia. Int Psychogeriatr. Apr 2014;26(4):555-563. [doi: 10.1017/S1041610213002329] [Medline: 24423470]
- Wang BR, Zheng HF, Xu C, Sun Y, Zhang YD, Shi JQ. Comparative diagnostic accuracy of ACE-III and MoCA for detecting mild cognitive impairment. Neuropsychiatr Dis Treat. 2019;15:2647-2653. [doi: 10.2147/NDT.S212328]
   [Medline: 31571881]
- 30. Beishon LC, Batterham AP, Quinn TJ, et al. Addenbrooke's Cognitive Examination III (ACE-III) and mini-ACE for the detection of dementia and mild cognitive impairment. Cochrane Database Syst Rev. Dec 17, 2019;12(12):CD013282. [doi: 10.1002/14651858.CD013282.pub2] [Medline: 31846066]
- 31. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. Apr 2005;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
- 32. Charernboon T. Diagnostic accuracy of the overlapping infinity loops, wire cube, and clock drawing tests for cognitive impairment in mild cognitive impairment and dementia. Int J Alzheimers Dis. 2017;2017:5289239. [doi: 10.1155/2017/5289239] [Medline: 28255496]
- 33. Howard C. P.012 SketchNet: equipping cognitive examinations with neural network computer vision. Can J Neurol Sci. Nov 2021;48(s3):S23. [doi: 10.1017/cjn.2021.294]
- 34. Obuchowski NA, Graham RJ, Baker ME, Powell KA. Ten criteria for effective screening: their application to multislice CT screening for pulmonary and colorectal cancers. AJR Am J Roentgenol. Jun 2001;176(6):1357-1362. [doi: 10.2214/ajr.176.6.1761357] [Medline: 11373191]

- 35. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. Inhal Toxicol. Nov 2014;26(13):811-828. [doi: 10.3109/08958378.2014.955932] [Medline: 25264934]
- 36. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association; 2013. [doi: 10.1176/appi. books.9780890425596]
- 37. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiol. Apr 1982;143(1):29-36. [doi: 10.1148/radiology.143.1.7063747] [Medline: 7063747]
- 38. Howard CW. Neural networks for cognitive testing: neural networks for cognitive test drawing classification and a new form of multiclass classification data visualization. Intelligence-Based Med. 2023;8(4). [doi: 10.1016/j.ibmed.2023. 100104]
- 39. Iandola FN. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv. Preprint posted online on Nov 4, 2016. [doi: 10.48550/ARXIV.1602.07360]
- 40. Youden WJ. Index for rating diagnostic tests. Cancer. Jan 1950;3(1):32-35. [doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3] [Medline: 15405679]
- 41. Efron B. Bootstrap methods: another look at the jackknife. Ann Statist. 1979;7(1). [doi: 10.1214/aos/1176344552]
- 42. Deng N, Allison JJ, Fang HJ, Ash AS, Ware JE. Using the bootstrap to establish statistical significance for relative validity comparisons among patient-reported outcome measures. Health Qual Life Outcomes. May 31, 2013;11:89. [doi: 10.1186/1477-7525-11-89] [Medline: 23721463]
- 43. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. Mar 2020;17(3):261-272. [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
- 44. Pedregosa F, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825-2830. URL: <a href="https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf">https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf</a> [Accessed 2025-06-04]
- 45. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with python. Presented at: Python in Science Conference; Jun 28 to Jul 3, 2010; Austin, TX. [doi: 10.25080/Majora-92bf1922-011]
- 46. Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. Sci Rep. Mar 13, 2024;14(1):6086. [doi: 10.1038/s41598-024-56706-x] [Medline: 38480847]
- 47. Bruno D, Schurmann Vignaga S. Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. Neuropsychiatr Dis Treat. 2019;15:441-447. [doi: 10.2147/NDT.S151253] [Medline: 30858702]
- 48. Calderón C, Beyle C, Véliz-García O, Bekios-Calfa J. Psychometric properties of Addenbrooke's Cognitive Examination III (ACE-III): an item response theory approach. PLOS One. 2021;16(5):e0251137. [doi: 10.1371/journal.pone.0251137] [Medline: 33956900]
- 49. Davis DH, Creavin ST, Yip JL, Noel-Storr AH, Brayne C, Cullum S. Montreal Cognitive Assessment for the detection of dementia. Cochrane Database Syst Rev. Jul 13, 2021;7(7):CD010775. [doi: 10.1002/14651858.CD010775.pub3] [Medline: 34255351]
- 50. Monaghan TF, Rahman SN, Agudelo CW, et al. Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value. Medicina (Kaunas). May 16, 2021;57(5):503. [doi: 10.3390/medicina57050503] [Medline: 34065637]
- 51. Berron D, Glanz W, Clark L, et al. A remote digital memory composite to detect cognitive impairment in memory clinic samples in unsupervised settings using mobile devices. NPJ Digit Med. Mar 26, 2024;7(1):79. [doi: 10.1038/s41746-024-00999-9] [Medline: 38532080]
- 52. Vyshedskiy A, Netson R, Fridberg E, et al. Boston cognitive assessment (BOCA) a comprehensive self-administered smartphone- and computer-based at-home test for longitudinal tracking of cognitive performance. BMC Neurol. Mar 15, 2022;22(1):92. [doi: 10.1186/s12883-022-02620-6] [Medline: 35291958]
- 53. Liu Z, Paek EJ, Yoon SO, Casenhiser D, Zhou W, Zhao X. Detecting Alzheimer's disease using natural language processing of referential communication task transcripts. J Alzheimers Dis. 2022;86(3):1385-1398. [doi: 10.3233/JAD-215137] [Medline: 35213368]
- 54. Howard C, Ng M. The autonomous cognitive examination: preliminary clinical trial results. J Neurol Sci. Dec 2023;455:121402. [doi: 10.1016/j.jns.2023.121402]
- 55. Howard C. Towards machine learning-based cognitive examination (S2.010). Neurology (ECronicon). May 3, 2022;98(18\_supplement):3914. [doi: 10.1212/WNL.98.18\_supplement.3914]

#### **Abbreviations**

**ACE-3:** Addenbrooke's Cognitive Examination-3

**AUC:** area under the curve

**DCA:** digital cognitive assessment **MoCA:** Montreal Cognitive Assessment

#### **RoCA:** Rapid Online Cognitive Assessment

Edited by Tobias Loetscher; peer-reviewed by Delphine Van Laethem, Anonymous, Wen Kwang Lim; submitted 21.09.2024; final revised version received 11.03.2025; accepted 12.03.2025; published 19.06.2025

Please cite as:

Howard C, Johnson A, Peedicail J, Ng MC

The Rapid Online Cognitive Assessment for the Detection of Neurocognitive Disorder: Open-Label Study

J Med Internet Res 2025;27:e66735

URL: https://www.jmir.org/2025/1/e66735

doi: 10.2196/66735

© Calvin Howard, Amy Johnson, Joseph Peedicail, Marcus C Ng. Originally published in the Journal of Medical Internet Research (<a href="https://www.jmir.org">https://www.jmir.org</a>), 19.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <a href="https://www.jmir.org/">https://www.jmir.org/</a>, as well as this copyright and license information must be included.