An Initial Exploratory Examination of the Sensitivity and Specificity of MyCog Mobile Using the Mini-Cog as a Proxy for Suspected Cognitive Impairment

Stephanie Ruth Young[1], Elizabeth Dworak[1], Greg Byrne[1], Yusuke Shono[2], Manrui Zhang[1] (iD), Julia Yoshino Benavente[1], Lihua Yao[1], Mike Bass[1] (iD), Laura Curtis[1], Maria Varela Diaz[1], Callie Jones[1], Richard Gershon[1], Michael Wolf[1], and Cindy Nowinski[1]

Abstract

Objectives: To help promote early detection of cognitive impairment in primary care, MyCog Mobile was designed as a cognitive screener that can be self-administered remotely on a personal smartphone. We explore the potential utility of MyCog Mobile in primary care by comparing MyCog Mobile to a commonly used screener, Mini-Cog. **Methods:** A sample of 200 older adults 65+ years (mean age = 72.56 years), completed the Mini-Cog and MyCog Mobile, which includes 2 memory measures and 2 executive functioning measures. A logistic regression model was conducted to predict failing Mini-Cog scores (≤ 2) based on MyCog Mobile measures. **Results:** A total of 20 participants earned a Mini-Cog score ≤ 2. MyCog Mobile demonstrated an AUC of 0.83 (95% bootstrap CI [0.75, 0.95]), sensitivity of 0.76 (95% bootstrap CI [0.63, 0.97]), and specificity of .88 (95% bootstrap CI [0.63, 0.10]). The subtest Name Matching from MyFaces and MySorting were the only significant predictors of failed Mini-Cogs. **Conclusions:** MyCog Mobile demonstrated sensitivity and specificity to identify participants who failed the Mini-Cog, and may show promise as a screening tool for cognitive impairment in older adults. Further research is necessary to establish the clinical utility of MyCog Mobile in a larger sample using documented clinical diagnoses.

Keywords

cognitive screening, primary care, app-based assessment, cognitive impairment, older adults

Dates received: 4 September 2024; revised: 18 October 2024; accepted: 21 October 2024.

Introduction

Primary care clinics are an ideal setting to identify cognitive impairment in the early stages, particularly since Medicare has covered cognitive screening as part of an Annual Wellness Visit since 2011.[1] However, primary care clinics face significant time and resource constraints, and thus cognitive screening measures must be fast, accurate, and easy to administer.[2] The Mini-Cog is one of the most commonly used screeners in primary care settings, as it is relatively easy to administer, and it takes about 3 min to complete.[3]

The Mini-Cog has demonstrated utility in community primary care settings. The brief screener is is relatively robust to education, gender, and ethnicity influences,[4]

and demonstrates an estimated 73% sensitivity and 84% specificity in detecting either MCI, dementia, or CI.[5] The Mini-Cog consists of 2 parts: a 3-word list recall (scored 1 point per word), and a clock drawing test (scored 0 or 2). Standard practice guidelines suggest participants who score 2 points or less on the Mini-Cog should be flagged for impairment, and further evaluation is recommended.[5]

[1]Northwestern University, Chicago, IL, USA
[2]Claremont Colleges, Claremont, CA, USA

Corresponding Author:
Stephanie Young, Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, 625 North Michigan Avenue, Suite 2700, Chicago, IL 60657, USA.
Email: stephanieruth.young@northwestern.edu

Despite the widespread use of the Mini-Cog, there are some disadvantages to using this tool in primary care. The Mini-Cog requires a trained examiner to administer following standardized instructions, which takes up valuable staff resources and time during ultra brief in-person visits. Moreover, although relatively simple to administer, human errors are common; 1 study found 61% of nurses made an error in the administration or scoring of the Mini-Cog.[6] The Mini-Cog is also recorded and scored by hand, and results must be entered into an Electronic Health Record (EHR), creating further steps in the workflow and opportunities for error.

MyCog Mobile is an app-based screener that may help overcome some of the limitations of the Mini-Cog. MyCog Mobile is self-administered on a patient's personal smartphone remotely prior to a primary care visit, and results are automatically sent into patient's EHR triggering appropriate best practice alerts.[7] The MyCog Mobile model saves valuable clinic time and staff resources and eliminates human errors in administration and scoring. While multiple digital screeners designed for at-home use are emerging on the market, few have been psychometrically validated, which limits their utility in clinical settings.[8] MyCog Mobile is unique in that it has demonstrated evidence of usability, feasibility, reliability, and validity in several previous studies.[7,9,10] However, it remains unclear how MyCog Mobile relates to well-established primary care screeners like the Mini-Cog. In this paper, we evaluated the sensitivity, specificity, and area under the curve of the MyCog Mobile tests using a Mini-Cog score 2 or less as a proxy for those who would be suspected of impairment and recommended for further evaluation under usual care standards for primary care .

Methods

A sample of 200 older adults was recruited by a market research agency through their large, nationally representative database of research volunteers. All participants were compensated and provided informed consent prior to participation. Inclusion criteria required participants to be age 65 years or older, speak English and provide informed consent prior to participation. Testing was conducted in 5 study facilities in Georgia, Texas, New Jersey, Florida, and Arizona. During the in-person visit, Mini-Cogs were administered first by a trained examiner to build rapport with participants before they started a completely self-administered smartphone task. Participants then self-administered MyCog Mobile on iPhones provided at the study facility. The MyCog Mobile battery takes approximately 20 min to complete. Given the brevity of both tests and the distinct task paradigms, we did not have significant concerns about the effects of counterbalancing. Following administration, Mini-Cogs were double-scored by 2 independent study staff, and any discrepancies were resolved with the study team to ensure consensus.

Measures

MyCog Mobile. *MyFaces* was originally developed by Rentz and colleagues to predict cerebral amyloid beta burden. This associative memory measure was adapted from the Mobile Toolbox Faces and Names measures.[11] Participants are shown 12 photos of faces paired with their names. After a 5- to 10-min delay, they complete 3 subtests: Recognition (choosing the correct face from 3 options), First Letter (indicating the first letter of the name of the displayed face), and Name Matching (selecting the correct name from 3 options). Each subtest generates a raw accuracy score.

MyCog Mobile. *MySorting* was adapted from the MyCog Dimensional Change Card Sorting and Mobile Toolbox Shape-Color Sorting tests.[11,12] MySorting reflects executive function and cognitive flexibility. Participants sort images by shape or color, as indicated by a cue word on the screen. Accuracy scores are based on the number of correct responses and response speed.

MyCog Mobile. *MyPictures* was adapted from MyCog Picture Sequence Memory and the Mobile Toolbox Arranging Pictures task.[11,12] MyPictures reflects episodic memory and involves arranging 16 images in a specific sequence. Participants first view the images in order and then recall their positions after the images are scrambled. Scores are based on exact matches and correctly ordered adjacent pairs. Exact match scores were generated for Trial 1 and Trial 2.

MyCog Mobile. *MySequence* is a working memory test adapted from the Mobile Toolbox Sequences test.[11] Mysequence involves remembering and ordering strings of letters and numbers. Participants are asked to recall the letters first, in alphabetical order, then the numbers, in ascending order. Trial difficulty increases until participants score incorrectly on 2 consecutive trials of the same length. Raw scores reflect the number of correct trials.

Analysis

All analyses were conducted in R statistical software.[13] With 20 cases (Mini-Cog ≤2) and 180 controls, our sample was 85% powered to detect an area under the curve (AUC) of 0.7 at the .05 significance level. We standardized all measure scores to place them on the same scale. We then conducted a logistic regression with the binary outcome based on Mini-Cog scores, with a score of 2 or lower indicating suspected cognitive impairment. The model predictors included scores from MyPictures Trial 1, MyPictures Trial 2, MyFaces First Letter, MyFaces Name Matching, MyFaces Recognition, MySequences, and MySorting. We did not use a composite score, but rather each individual variable as predictors in the model.

We used pseudo R^2 to evaluate the acceptability of the model fit, with values of .2 or higher representing excellent model fit.[14] We considered coefficients for the predictors significant at the .05 level. Our primary performance metrics were AUC, sensitivity, and specificity. Generally, a model with an AUC of 0.7 to 0.8 is considered acceptable, and 0.8 to 0.9 is considered excellent in its ability to discriminate between cases and controls (e.g., suspected impairment versus no impairment suspected).[15] As an additional metric of test utility, we considered a combined sum of sensitivity and specificity of 1.5 or greater to be acceptable.[16]

Given the small number of cases in our sample, we generated 1000 bootstrap samples from the complete dataset and fitted a logistic regression model to each simulated sample with replacement. Bootstrap confidence intervals for AUC, sensitivity, and specificity were computed using percentile methods to provide a 95% confidence range for each metric. To mitigate potential overfitting, we also ran an optimism-corrected ROC to address potential bias of the model fit.[17,18] Using bootstrapping, this method adjusts the ROC AUC to provide a more accurate estimate of the model's generalizability on new data, thereby offering a more reliable estimate of the model's predictive ability.

Results

A sample of 200 older adults ages 65 to 87 years (mean age = 72.56 years; SD = 5.11 years) were enrolled in this study (Table 1). A total of 20 participants earned a Mini-Cog score of 2 or less.

The AUC of MyCog Mobile was 0.83 (bootstrapped 95% CI: [0.75, 0.95]), with a sensitivity of 0.76 (bootstrapped 95% CI: [0.63, 0.97]) and specificity of 0.88 (bootstrapped 95% CI: [0.63, 0.10]). The pseduo R^2 value of .25 indicated an excellent logistic regression model fit to the data. MyFaces Name Matching and MySorting were the only significant predictors of failed Mini-Cogs (see Table 2).

The model initially achieved an AUC of 0.83 when fitted to the original dataset. After applying optimism-correction, the ROC AUC decreased to 0.77. This reduction indicates that while the initial AUC value may have been slightly overestimated, it fell within the bootstrapped confidence interval and above acceptable a priori thresholds. The distribution of AUC values for the bootstrap samples evaluated on the original dataset, and the estimated apparent and adjusted test performance are illustrated in Figure 1.

Discussion and Limitations

Findings from this exploratory study suggest that MyCog Mobile demonstrates adequate sensitivity and specificity to identify participants who fail a Mini-Cog. MyFaces Name Matching, a memory subtest, and MySorting, an executive functioning test, were the best predictors of a failed Mini-Cog in our model. These results should be considered preliminary for several reasons. It is important to stress we used a general population sample and did not have information about the participants clinical diagnostic status. Although the Mini-Cog is commonly used to flag individuals for further evaluation, it is not a diagnostic tool on its own nor does it have perfect accuracy to detect impairment. Thus, it is unclear if any participants in our study would be
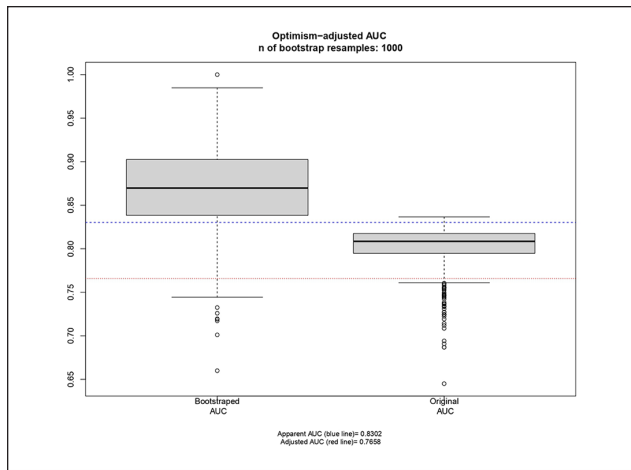
Table 1. Sample Demographics.

	n (%)
Total N	200 (100)
Gender	
Female	109 (45.5)
Male	91 (54.5)
Racial identity	
Asian	1 (0.5)
Black or African American	35 (17.5)
Middle Eastern or North African	1 (0.5)
Native Hawaiian or Pacific Islander	1 (0.5)
White or European American	157 (78.5)
Other identity	8 (4)
Prefer not to answer	1 (0.5)
Ethnic identity	
Hispanic	40 (20)
Not Hispanic	160 (80)
Education	
Less than high school	1 (0.5)
High School Diploma or GED	38 (19)
Some college or 2-year degree	65 (32.5)
4-year college degree	61 (30.5)
Graduate degree	35 (17.5)
Mini-cog score	
0	1 (0.5)
1	5 (2.5)
2	14 (7)
3	31 (15.5)
4	49 (24.5)
5	100 (50)

Table 2. Logistic Regression Predictors of Suspected Cognitive Impairment Indicated by Mini-Cog.

	β (SE)	*P* value
MyPictures		
Trial 1	0.24 (0.44)	.59
Trial 2	0.47 (0.39)	.23
MyFaces		
First letter	−0.44 (0.40)	.26
Recognition	0.24 (0.31)	.44
Name matching	0.78 (0.35)	.02*
MySequences	0.22 (0.29)	.44
MySorting	0.79 (0.32)	.01*

*P < .05.

Figure 1. The optimism-adjusted AUC (0.77) is lower than the unadjusted AUC (0.83) but above acceptable a priori threshold of 0.70.

diagnosed with a cognitive impairment in a clinical setting. An estimated 17% of community-dwelling older adults live with MCI,[19] and the ratio of suspected cases to controls in our sample (20:180) was small for analysis purposes, which increases the risk of bias in our findings. We mitigated this risk by applying an optimism correction to our bootstrap resampled estimates, which suggested that the AUC in a sample of independent data would likely be acceptable, yet this will need to be confirmed in future work.

The generalizability of these results is further limited by lack of representation of racial identities other than Black/ African American or White/European American, and few participants with less than a high school education. Moreover, participants self-administered MyCog Mobile in a research setting, not in their own homes prior to a primary care clinic visit. It will be important to examine the feasibility of completing MyCog Mobile in a home setting, taking into consideration the technology experience, effort, and time required of the patients to complete MyCog Mobile, especially among vulnerable subgroups with worse physical and mental health. Findings support clinical validation of MyCog Mobile in a sample that includes healthy older adults as well as those with confirmed diagnoses of cognitive impairment in primary care settings.

Conclusions

MyCog Mobile shows promise as a potential screening tool for cognitive impairment in older adults. While preliminary, these findings suggest MyCog Mobile demonstrates adequate sensitivity and specificity to detect a failed Mini-Cog, with the MyFaces and MySorting subtests showing the most potential for detection. Further validation in a clinical setting with a larger sample, including individuals with

confirmed cognitive diagnoses, is necessary to establish the clinical utility of MyCog Mobile and determine which subtests should be prioritized in the battery.

Ethical Approval

This study is determined to be non-human subject research by Northwestern University IRB (study #STU00221000) on February 22, 2024.

Consent to Participant

All participants provided written informed consent before participating.

Consent to Publication

Not applicable.

ORCID iDs

Manrui Zhang (iD) https://orcid.org/0000-0001-8471-8578

Mike Bass (iD) https://orcid.org/0000-0002-7732-6557

Data Availability Statement

Data will be shared in a public data repository upon completion of all grant-related activities in 2026.

References

1. Jacobson M, Thunell J, Zissimopoulos J. Cognitive assessment at medicare's annual wellness visit in fee-for-service and medicare advantage plans: study examines the use of medicare's annual wellness visit and receipt of cognitive assessment among medicare beneficiaries enrolled in fee-for-service medicare or medicare advantage. *Health Affairs*. 2020;39(11):1935-1942.
2. Liss J, 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*. 2021;290(2):310-334.
3. Scott J, Mayo AM. Instruments for detection and screening of cognitive impairment for older adults in primary care settings: a review. *Geriatr Nurs*. 2018;39(3):323-329.

4. Milne A, Culverwell A, Guss R, Tuppen J, Whelton R. Screening for dementia in primary care: a review of the use, efficacy and quality of measures. *Int Psychogeriatr*. 2008; 20(5):911-926.

5. Abayomi SN, Sritharan P, Yan E, et al. The diagnostic accuracy of the Mini-Cog screening tool for the detection of cognitive impairment: a systematic review and meta-analysis. *PLoS One*. 2024;19(3):e0298686.

6. Tam E, Gandesbery BT, Young L, Borson S, Gorodeski EZ. Graphical instructions for administration and scoring the Mini-Cog: results of a randomized clinical trial. *J Am Geriatr Soc*. 2018;66(5):987-991.

7. Young SR, Dworak EM, Byrne GJ, et al. Remote self-administration of cognitive screeners for older adults prior to a primary care visit: pilot cross-sectional study of the Reliability and usability of the Mycog mobile screening App. *JMIR Form Res*. 2024;8(1):e54299.

8. Sabbagh MN, Boada M, Borson S, et al. Early detection of mild cognitive impairment (MCI) in an at-home setting. *J Prevent Alzheimer's Dis*. 2020;7:171-178.

9. Young SR, Lattie EG, Berry AB, et al. Remote cognitive screening of healthy older adults for primary care with the Mycog mobile App: iterative design and usability evaluation. *JMIR Form Res*. 2023;7:e42416.

10. Young SR, Dworak EM, Byrne GJ, et al. Protocol for a construct and clinical validation study of MyCog Mobile: a remote smartphone-based cognitive screener for older adults. *BMJ Open*. 2024;14(4):e083612.

11. Nowinski CJ, Kaat A, Slotkin J, et al. 2 validity and reliability of mobile toolbox cognitive assessments. *J Int Neuropsychol Soc*. 2023;29(s1):780-781.

12. Hosseinian Z, Curtis L, Shono Y, et al. Validation of MyCog: a self-administered rapid detection tool for cognitive impairment in everyday clinical settings (P14-3.001). *Neurology*. 2022;98(18_supplement):3434.

13. Team RDC. *R: A Language and Environment for Statistical Computing*. R Foundaton for Statistical Computing, 2010.

14. McFadden D. Quantitative methods for analysing travel behaviour of individuals: some recent developments. In: Hensher DA, Stopher PR, eds. *Behavioural Travel Modelling*. Routledge; 2021:279-318.

15. Hosmer Jr DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. John Wiley & Sons; 2013.

16. Power M, Fell G, Wright M. Principles for high-quality, high-value testing. *BMJ Evid-Based Med*. 2013;18(1):5-10.

17. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans M, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781.

18. Smith GC, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. *Am J Epidemiol*. 2014;180(3):318-324.

19. Pessoa RMP, Bomfim AJL, Ferreira BLC, Chagas MHN. Diagnostic criteria and prevalence of mild cognitive impairment in older adults living in the community: a systematic review and meta-analysis. *Arch Clin Psychiatry*. 2019;46(3):72-79.