AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Assessing the Utility of the Montreal Cognitive Assessment in Screening for Cognitive Impairment in Patients With Systemic Lupus Erythematosus

Oshrat E. Tayer-Shifman,[1] Kimberley Yuen,[2] Robin Green,[3] Mahta Kakvan,[4] Patricia Katz,[5]
Kathleen S. Bingham,[6] Juan Pablo Diaz-Martinez,[4] Lesley Ruttan,[3] Joan E. Wither,[7] Maria Carmela Tartaglia,[8]
Jiandong Su,[4] Dennisse Bonilla,[4] May Y. Choi,[9] Simone Appenzeller,[10] Michelle Barraclough,[4]
Dorcas E. Beaton,[11] and Zahi Touma[12]

Objective. Screening for cognitive impairment (CI) in systemic lupus erythematosus (SLE) relies on the American College of Rheumatology (ACR) neuropsychological battery (NB). By studying the concurrent criterion validity, our goal was to assess the Montreal Cognitive Assessment (MoCA) as a screening tool for CI compared to the ACR-NB and to evaluate the added value of the MoCA to the Automated Neuropsychological Assessment Metrics (ANAM).

Methods. A total of 285 adult SLE patients were administered the ACR-NB, MoCA, and ANAM. For the ACR-NB, patients were classified as having CI if there was a Z score of ≤-1.5 in ≥2 domains. The area under the curve (AUC) and sensitivities/specificities were determined. A discriminant function analysis was applied to assess the ability of the MoCA to differentiate between CI, undetermined CI, and non-CI patients.

Results. CI was not accurately identified by the MoCA compared to the ACR-NB (AUC of 0.66). Sensitivity and specificity were poor at 50% and 69%, respectively, for the cutoff of 26, and 80% and 45%, respectively, for the cutoff of 28. The MoCA had a low ability to identify CI status. The addition of the MoCA to the ANAM led to improvement on the AUC by only 2.5%.

Conclusion. The MoCA does not have adequate concurrent criterion validity to accurately identify CI in patients with SLE. The low specificity of the MoCA may lead to overdiagnosis and concern among patients. Adding the MoCA to the ANAM does not substantially improve the accuracy of the ANAM. These results do not support using the MoCA as a screening tool for CI in patients with SLE.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a multisystem chronic autoimmune disease (1) in which the nervous system is commonly affected (2). There are 19 neuropsychiatric SLE syndromes defined by the American College of Rheumatology (ACR) (3,4). Cognitive impairment (CI) is one of the most common manifestations of neuropsychiatric SLE, with a prevalence ranging from 20% to 80% (5–7) and a pooled prevalence of 38% (5). CI has a significant effect on patients' health-related quality of life

---

SIGNIFICANCE & INNOVATIONS

- This study is the largest to date evaluating the utility of the Montreal Cognitive Assessment (MoCA) as a screening tool for cognitive impairment (CI) in patients with systemic lupus erythematosus (SLE).
- The MoCA does not have adequate concurrent criterion validity compared to the American College of Rheumatology neuropsychological battery (gold standard) for the screening for CI in SLE patients.
- Due to the low ability of the MoCA to identify the cognitive status of SLE patients, this study cannot support the MoCA as a screening tool for CI in patients with SLE.
- The addition of the MoCA to the Automated Neuropsychological Assessment Metrics (ANAM) does not substantially improve the accuracy of the ANAM.

---

and participation in social, leisure, work, and family-related activities [8–13].

The ACR neuropsychological battery (ACR-NB) is the suggested validated tool for assessing cognitive function in adult SLE patients [3,14]. The ACR-NB requires ~1 hour of specialized personnel time to administer, plus additional time for scoring and interpretation. For many clinics, these are notable barriers to cognitive function assessment [15,16]. Recently, we have described evidence for the validity of the Automated Neuropsychological Assessment Metrics (ANAM) as a screening tool for CI in patients with SLE. The ANAM is a computerized battery that takes 20 to 40 minutes to complete, depending on the subtests used, and does not require a specialist to administer. We have shown that CI could be accurately identified by selected ANAM subtests and scores, with the best models demonstrating excellent area under the curve (AUC) values of 81% and 84% [17]. We have described 2 different approaches to interpret ANAM data in SLE. The first approach uses a total score index (ANAM index) with different cutoffs generated from regression analyses [17]. The second approach is based on a simple decision tree generated from classification and regression tree analysis [18]. The Montreal Cognitive Assessment (MoCA) is a brief paper-and-pencil assessment tool designed for the screening of mild neurocognitive disorder (previously known as mild cognitive impairment) in the general population [19]. It is freely accessible for clinical and educational purposes and is available in nearly 100 languages. One-hour online training and certification has been mandatory for its use since September 2019. The instrument provides a cursory assessment of memory, visuospatial skills, executive functions, attention and concentration, language, and orientation, and it takes 10 minutes to administer.

The aims of this study were to asses the concurrent criterion validity of the MoCA to evaluate its utility as a screening tool for CI in SLE patients compared to the ACR-NB, and to determine the added value of the MoCA to the ANAM. We hypothesized that the MoCA will perform adequately in screening for CI in patients with SLE.

PATIENTS AND METHODS

Study population. A total of 285 adult SLE patients who attended the University of Toronto Lupus Clinic between July 2016 and March 2020 participated in the study. Inclusion criteria were: 1) fulfillment of the 1997 revised ACR criteria for SLE classification or 3 criteria and a supportive biopsy [1,20]; 2) age 18–65 years; and 3) the ability to give informed consent. Exclusion criteria were mental or physical disability preventing participation in the study, and semifluency in English or less, precluding valid completion of verbal items of the ACR-NB. Of the 838 screened patients, 778 were eligible for participation and 412 patients provided informed consent. Of the 412 patients, 49 withdrew from the study (citing duration and stress of visits) and 285 patients actively participated (78 have yet to participate). This project was approved by the University Health Network Research Ethics Board.

Procedures and outcome measures. All patients were administered the MoCA first, followed by the ACR-NB and the ANAM on the same day. Patients' scores on the ACR-NB were compared to normative data to obtain Z scores [17].

The original ACR battery [21] consists of 11 cognitive tests representative of 6 cognitive domains (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971). Our ACR-NB is identical to the ACR recommended cognitive battery for adults with SLE [14,21], except that the Hopkins Verbal Learning Test–Revised (HVLT-R) was substituted for the California Verbal Learning Test [22]. A domain was defined as impaired if a Z score of ≤–1.5 was reached in at least 1 test in the following domains: manual motor speed, simple attention and processing speed, visual-spatial construction, and language processing, or a Z score of ≤–1.5 in 2 or more tests in the following domains: learning, memory, and executive functioning [17]. We also conducted sensitivity analyses using different definitions for CI; first CI was defined as ≥2 domains with a Z score of ≤–2, and second CI was defined as ≥1 domain with a Z of score ≤–2.

The ANAM (version 4) General Neuropsychological Screening battery consists of 15 subtests (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971). For each ANAM subtest, several scores are provided: percentage of correct responses, mean reaction time, throughput, and coefficient of variation of reaction time, as detailed elsewhere [17].

The MoCA is a 1-page, 10-minute paper-and-pencil assessment tool designed for screening of mild neurocognitive disorder in the general population ages 55–85 years [19]. The questionnaire is accessible at https://www.mocatest.org/ and includes

12 tasks with a maximum score of 30 points, assessing memory, visuospatial skills, executive functions, attention and concentration, language, and orientation (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971), grouped into 7 categories: visuospatial/executive, naming, attention, language, abstraction, delayed recall, and orientation. A score of 26 is recommended as the threshold, below which a patient is classified as having CI (mild or greater) in the general population. We compared the performance of the MoCA to the ACR-NB (gold standard). For the ACR-NB, patients were classified on the ACR-NB as CI if there was a Z score of ≤–1.5 in ≥2 domains, non-CI if no domain had a Z score of ≤–1.5, or indeterminate if there was a Z score of ≤–1.5 in only 1 domain.

Statistical analysis. Demographic and clinical characteristics were summarized as the percentage of patients, mean ± SD, and median and interquartile range. The performance of the MoCA in the identification of CI as classified by the ACR-NB (criterion measure) was evaluated using receiver operating characteristic (ROC) curve analysis and AUC calculation. Values of the AUC can be interpreted as outstanding (1.0–0.91), excellent (0.81–0.90), good (0.71–0.8), fair (0.61–0.7), and poor (≤0.6) performance for identifying CI [23]. Descriptive statistics and 2-by-2 contingency tables with different MoCA cutoffs (24–28) were applied to determine their sensitivity and specificity to CI.

Diagnostic test properties, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, and negative likelihood ratio were determined (with 95% confidence intervals). A sensitivity and specificity ≥75% were proposed to demonstrate concurrent criterion validity evidence for MoCA with the ACR-NB [24]. A likelihood ratio >10 provides strong evidence to rule in CI and below 0.1 to rule out CI, while 1 suggests little value toward a diagnosis [25,26]. Cohen's kappa coefficient values were derived to assess agreement between the ACR-NB and MoCA. Levels of agreement can be interpreted as poor (≤0.20), fair (0.21–0.40), moderate (0.41–0.60), and good (≥0.61) [27]. Additionally, a discriminant function analysis was applied to assess the ability of the MoCA to differentiate between CI, undetermined CI (1 domain with a Z score of ≤–1.5), and non-CI patients as classified by the ACR-NB. We also analyzed the data after adjusting for ethnicity and education in the discriminant function analysis. We did not adjust for age, since age was used in all Z score calculations in the ACR-NB. The scores from the 7 MoCA categories were compared between CI and non-CI patients classified by ACR-NB to assess whether the MoCA tasks can differentiate between CI and non-CI patients; hypotheses tests were performed by Student's *t*-test and effect size were calculated.

An additional analysis was conducted to determine whether the MoCA improves the performance of the ANAM to identify CI as compared to the ACR-NB (MoCA plus ANAM versus ACR-NB). For this analysis, the ANAM index score was standardized to obtain a Z score [17]. The ANAM index score was calculated for each patient according to its formula, using selected ANAM subtests and scores:

$$
\begin{aligned}
\text{ANAM index} = {}& 31.85 - 0.06 \times \text{PCT/CSD} - 0.14 \\
& \times \text{PCT/GNG} - 9.93 \times \text{CV/SP} - 6.38 \\
& \times \text{CV/TCRT} + 9.74 \times \text{MR/TL} - 0.06 \\
& \times \text{TP/CSL} - 0.02 \times \text{TP/SRTR} - 0.0008 \\
& \times \text{MS/TPZ}
\end{aligned}
$$

For details refer to Tayer-Shifman et al (2020) [17]: CSD = code substitution delay; CSL = code substitution learning; CV = coefficient of variation of reaction time; GNG = go/no go; MR = mean reaction time; PCT = percentage of correct responses; PRT = procedural reaction time; SP = spatial processing; SRTR = simple reaction time repeated; TCRT = 2-choice reaction time; TL = tapping left hand; TP = throughput; MS = mean-score; TPZ = tower puzzle [17].

MoCA scores were negated (multiplied by –1) to make them compatible with the ANAM index (i.e., a higher score corresponded to a higher probability of CI) and standardized to obtain Z scores. The Z scores of the ANAM index score and the MoCA were summed to create a combined ANAM-MoCA Z score, with a mean ± SD of 0 ± 2, and with equal weights from both instruments. The performance of the ANAM-MoCA Z score to accurately identify CI, as classified by the ACR-NB, was evaluated using ROC analysis. The AUC was calculated for ANAM, MoCA, and ANAM-MoCA Z score compared to the ACR-NB.
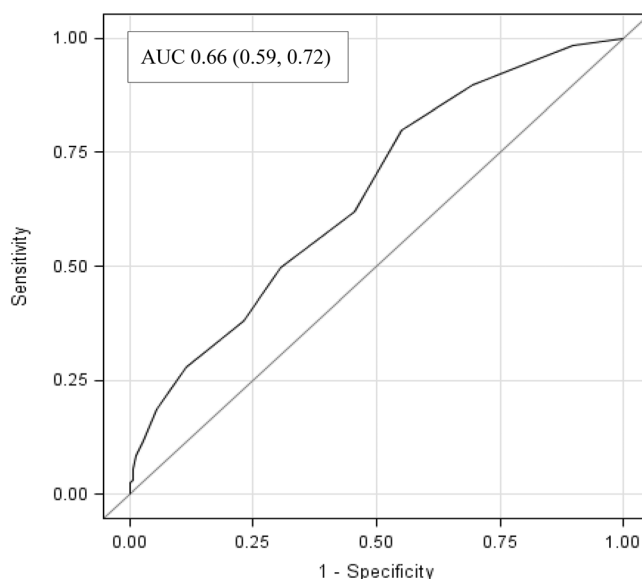

Figure 1. Receiver operating characteristic curves for identifying cognitive impairment based on the performance of the Montreal Cognitive Assessment compared to the American College of Rheumatology neuropsychological battery (n = 276). AUC = area under the curve (95% confidence interval).

Table 1. Performance of the Montreal Cognitive Assessment (MoCA) compared to the ACR-NB*

MoCA cutoff	AUC†	Sensitivity, %	Specificity, %	PPV, %	NPV, %	LR+	LR–
<24	0.66 (0.59–0.72)	28 (20–30)	88 (83–94)	68 (55–80)	58 (52–65)	2.41 (1.14–3.68)	0.82 (0.72–0/91)
<25	0.66 (0.59–0.72)	38 (30–46)	77 (70–84)	59 (48–70)	59 (52–66)	1.64 (1.04–2.25)	0.81 (0.68–0.94)
<26	0.66 (0.59–0.72)	50 (41–58)	69 (62–77)	59 (49–68)	61 (54–68)	1.62 (1.14–2.11)	0.73 (0.58–0.87)
<27	0.66 (0.59–0.72)	62 (54–70)	54 (46–62)	54 (46–62)	62 (54–70)	1.36 (1.06–1.66)	0.7 (0.51–0.88)
<28	0.66 (0.59–0.72)	80 (73–87)	45 (37–53)	56 (49–63)	72 (63–81)	1.45 (1.2–1.7)	0.45 (0.28–0.62)

* Values in parentheses are the 95% confidence interval. ACR-NB = American College of Rheumatology neuropsychological battery; AUC = area under the curve; LR+ = positive likelihood ratio; LR– = negative likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.
† AUC was calculated using continuous data and regression analysis to identify the best cutoff based on the Youden index.

RESULTS

MoCA versus ACR-NB. A total of 285 SLE patients were enrolled; 253 (88.8%) were female and the mean ± SD age and SLE disease duration at the study visit were 41.3 ± 12.1 years and 14.4 ± 10.1 years, respectively (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971). A total of 276 patients completed the MoCA and the ACR-NB. Based on the ACR-NB, 129 patients (47%) were classified as having CI, 85 (31%) as undetermined CI, and 62 (22%) as non-CI patients.

CI was not accurately predicted by MoCA, but had a fair AUC of 0.66 (95% confidence interval 0.59–0.72) (Figure 1 and Table 1). A cutoff of 26 yielded a sensitivity, specificity, PPV, and NPV of 50%, 69%, 59%, and 61%, respectively, with a positive likelihood ratio of 1.62 and negative likelihood ratio of 0.73. A MoCA cutoff of 28 yielded the highest accuracy (highest Youden index), with a sensitivity of 80% but with a clear tradeoff in specificity of 45%, and with PPV and NPV of 56% and 72%, respectively, with a positive

likelihood ratio of 1.45 and negative likelihood ratio of 0.45 (Table 1). The distribution of the MoCA scores in the cohort demonstrates the overlapping scores between CI and non-CI patients based on ACR-NB (Figure 2). In addition, the agreement (Cohen's kappa coefficient) between the ACR-NB and MoCA was poor to fair using different CI definitions by the ACR-NB: for CI defined as a Z score of ≤–1.5 in ≥2 domains, κ = 0.192 (95% confidence interval 0.08–0.31), for CI defined as ≥2 domains with a Z score of ≤–2. CI, κ = 0.249 (95% confidence interval 0.13–0.36), and for CI defined as ≥1 domain with a Z score of ≤–2, κ = 0.173 (95% confidence interval 0.08–0.26) (see Supplementary Table 3A, B, and C, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971).

Figure 3 shows the predicted probability of having CI based on the MoCA score. There was a 25% predicted probability of CI even with the highest score (28) on the MoCA. Using different CI definitions by the ACR-NB led to similar results and did not improve the performance of MoCA compared to the ACR-NB
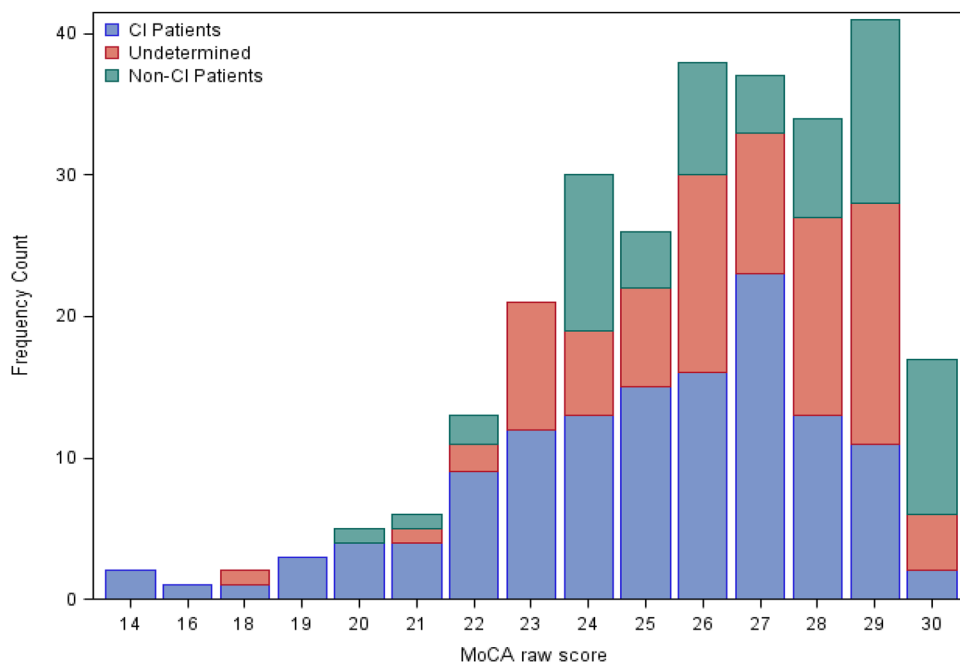

Figure 2. The distribution of Montreal Cognitive Assessment (MoCA) scores in cognitive impairment (CI) and non-CI patients (classified by the American College of Rheumatology neuropsychological battery).
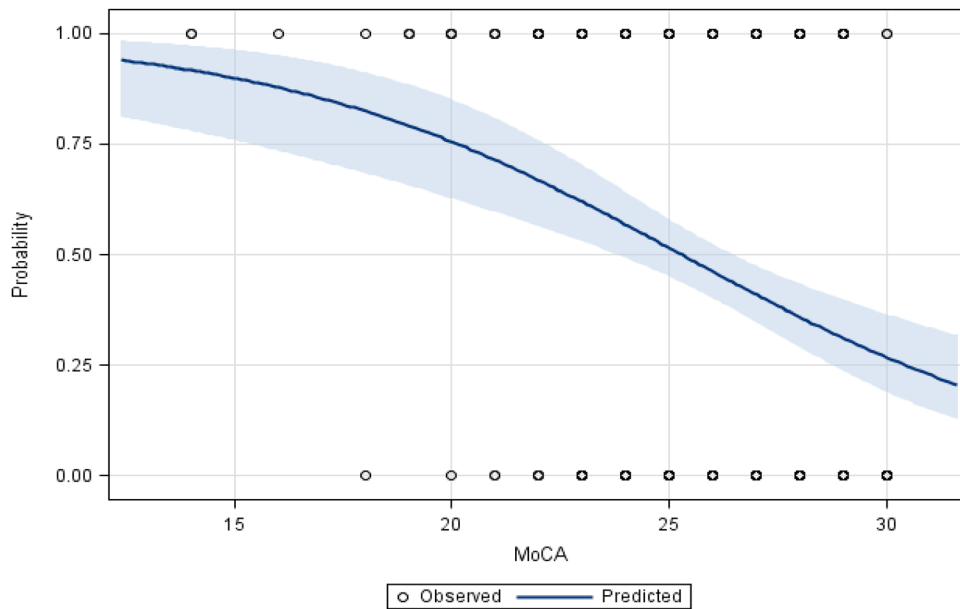
Figure 3. The predicted probability of having cognitive impairment (classified by the American College of Rheumatology neuropsychological battery) and 95% confidence intervals by Montreal Cognitive Assessment (MoCA) scores (n = 276).

(see Supplementary Table 4A and 4B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971).

Based on the discriminant function analysis, the misclassification rates in the 3 categories (CI, undetermined CI, and non-CI) were 50%, 83%, and 44%, respectively. All 3 categories overlapped between the scores 25 to 27 (cutoffs 26–28). Thus, the MoCA had a low ability to accurately identify CI status (Figure 4)

when compared to these more detailed approaches. Ethnicity did not change the results of this analysis.

Significant differences/effect sizes between the ACR-NB–classified CI and non-CI groups were seen in the following MoCA categories: delayed recall 0.63 ($P < 0.0001$), visuospatial/executive 0.47 ($P < 0.0001$), attention 0.43 ($P = 0.0003$), and naming 0.23 ($P = 0.04$). MoCA categories of abstraction, language, and orientation were not significantly different between CI and non-CI patients.

Figure 4. The probability of belonging to 3 cognitive impairment (CI) groups (classified by the American College of Rheumatology neuropsychological battery) at each Montreal Cognitive Asse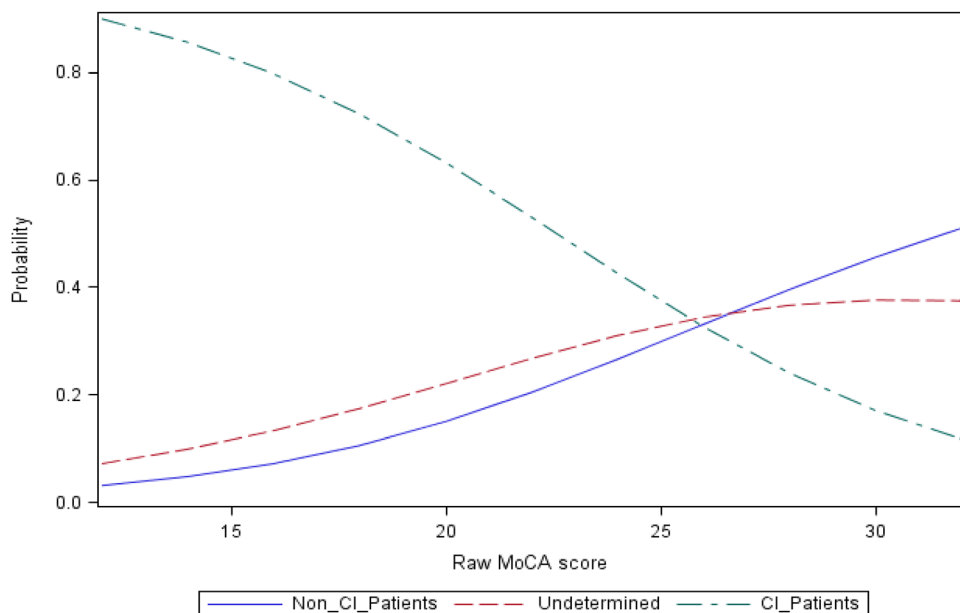ssment (MoCA) score (n = 276). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971/abstract.

Table 2. Studies evaluating Montreal Cognitive Assessment performance in systemic lupus erythematosus*

Author, year (ref.)	No. pts.	Anchor tool	Controls	CI prevalence, %†	AUC	Se/Sp, cutoff <26, %	Se/Sp, cutoff <28, %
Adhikari et al, 2011 (30)	44	ANAM	RA patients	25	NA	83/73	100/36
Nantes et al, 2017 (28)	98	HVLT-R	Normal population	31	0.71	73/63	
Paez-Venegas et al, 2019 (32)	44	ACR-NB	Normal population	70	0.99	84/100	
Chalhoub and Luggen, 2019 (31)	78	ANAM	Normal population and RA patients	39.7–80.8‡	NA	52–63/ 59–68‡	83–94/ 36–46‡
Raghunath et al, 2021 (37)	95	ACR-NB	Normal population	19–49‡	0.69–0.78‡	47/94	83/58
Papastefanakis et al, 2021 (39)	71	NB	Normal population and RA patients	9–15‡	NA	91/45	100/14
Current study	276	ACR-NB	Normal population	47	0.66	50/69	80/45

* ACR-NB = American College of Rheumatology neuropsychological battery; ANAM = Automated Neuropsychological Assessment Metrics; AUC = area under the curve; CI = cognitive impairment; HVLT-R = Hopkins Verbal Learning Test–Revised; NA = not available; NB = neuropsychological battery (any battery not similar to the ACR-NB); RA = rheumatoid arthritis; ref. = reference; Se = sensitivity; Sp = specificity.
† By the anchor tool used.
‡ Value range depends on CI definition.

MoCA plus ANAM versus ACR-NB. As previously described, the ANAM index showed a good ability to identify CI compared to the ACR-NB, with an AUC of 0.79 (95% confidence interval 0.71–0.88) (17). Adding the MoCA score to create the ANAM-MoCA Z score improved the AUC to 0.81 (95% confidence interval 0.73–0.89), which led to improvement in the sensitivity from 72% to 77%, with the same specificity of 78% at the highest Youden index (see Supplementary Figure 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971). Thus, the improvement in accuracy of CI identification for the combined ANAM-MoCA Z score compared to the ANAM index alone was only 2.5% over the ANAM alone.

DISCUSSION

This is the largest study to date evaluating the utility of the MoCA as a screening tool for CI in patients with SLE. We compared the performance of the MoCA to the ACR-NB, the gold standard for assessing cognitive function in patients with SLE. Additionally, we are the first to study the performance of the ANAM and MoCA together, attempting to enhance the accuracy of the former. Unfortunately, this study showed that the utility of the MoCA in screening for CI in SLE patients is low. The MoCA failed to reach an acceptable AUC compared to the ACR-NB as a gold standard. The standard MoCA cutoff of 26 for mild CI demonstrated inadequate performance, yielding poor sensitivity of only 50% and specificity of 69%. The low sensitivity may lead to underdiagnosis and an inability to identify patients who could benefit from early intervention for CI. Using higher cutoffs generated better sensitivities but had unacceptably poor specificities and a low ability to accurately identify CI status. The low specificity of the MoCA cutoff of 28 will lead to overdiagnosis and can cause unnecessary concern for the patients who screen positive on the MoCA. Moreover, an overdiagnosis will lead to more complex,

expensive testing (the ACR-NB). Furthermore, the addition of the MoCA to the ANAM, despite upgrading the AUC from the good to excellent category (from 79% to 81%), yielded an incremental improvement of only 2.5%.

We have recently shown in our systematic review that the assessment of CI in SLE remains heterogeneous (29). Moreover, there is no consistency between studies on the definitions of CI using the ACR-NB, and different studies have used various CI definitions (5). Definitions for CI have varied, with cutoffs ranging from 1 to 3 SDs below normative values on domains/tests. We chose to classify CI by the ACR-NB if Z scores were ≤–1.5 in 2 domains or more. Using different CI definitions by the ACR-NB led to similar results, with poor performance of the MoCA compared to the ACR-NB (see Supplementary Table 4A and 4B, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24971).

Several studies used the MoCA to assess cognitive function in SLE patients (13,28,30–36). The prevalence of CI using the MoCA in these studies varies between 29.5% and 67.9%. While differences may be due in part to demographic and clinical differences in the samples of these studies, we suggest that differences were also attributable to limitations of the tool itself. We found a prevalence of 39.5% by MoCA in this cohort using the standard cutoff of 26 compared to a prevalence of 46% using the ACR-NB. Although the prevalences are similar, the MoCA had high misclassification rates of 50% in classifying CI and 44% in classifying non-CI, leading to misclassification of 45 patients as CI and 65 patients as non-CI (while the ACR-NB classified them in the opposing groups).

Fewer studies examine the performance of the MoCA in assessing the cognitive function of SLE patients compared to other tools (Table 2). These studies also exhibited large discrepancies in the MoCA's performance, with sensitivities ranging from 47% to 84% and specificities from 63% to 100% using the

standard cutoff of 26. Again, differences in the sample as well as statistical methods and the use of different external anchors between studies would have played a role (5,28,31,32), but the disparity also is compatible with our contention regarding the modest validity of the MoCA for measuring CI in this population.

Only 2 previous studies directly compared the MoCA to the ACR-NB in patients with SLE. Paez-Venegas et al (32) showed an outstanding performance of the MoCA, with an AUC of 99.4 ($P < 0.001$), 84% sensitivity, and 100% specificity. However, these results have not been replicated in any of the subsequent studies and may have been influenced by atypically small sample sizes. Paez-Venegas et al studied 44 SLE patients, of which nearly 20% were men with active disease (mean ± SD SLEDAI score of 6.5 ± 4.6) and found a high prevalence of CI of 70%. In the study of Paez-Venegas et al, CI was defined when only 1 domain was impaired, which is not the usually accepted definition of CI (29). In a recent study, Raghunath et al (37) compared the performance of the MoCA to the ACR-NB, studying 95 SLE patients. Using an identical CI definition by the ACR-NB as ours, the accuracy of the MoCA with the standard cutoff of 26 was moderate at best, with an AUC of 0.70, just slightly higher than our AUC of 0.66, and with poor sensitivity of only 47%. A cutoff of 27 on the MOCA had the highest accuracy, with a good AUC of 0.77, but yielded a sensitivity of only 72% and a specificity of 83%. The necessity for enhanced sensitivity led the authors to recommend a higher MoCA cutoff of 28, with a sensitivity of 83% and this higher cutoff penalized the accuracy and specificity, at 0.71 and 58%, respectively. Adapting a cutoff of 28 to our study led to good sensitivity of 80% as well, but with a very low specificity of 45% and a high misclassification rate, as was shown by the discriminant function analysis. More importantly, Raghunath's study used binary outcomes for AUC calculations, which may have led to overestimation of the AUCs (38), while our study used continuous data and regression analysis to identify the best cutoff based on the Youden index.

Papastefanakis et al (39) compared the MoCA to a short NB comprising only 11 tests in 77 SLE patients. They used 3 CI definitions using a normal population or rheumatoid arthritis (RA) patients as controls. The CI definition using the normal population as controls was relatively stringent and required at least 2 SDs below the population average on at least 3 of 11 subtests. This approach classified only 11 patients as impaired, yielding a very low prevalence of CI of 15%. A MoCA cutoff of 26 yielded a sensitivity of 91% and poor specificity of 45%, probably due to a small sample size and stringent CI definition, leading to classifying only severely impaired patients as CI.

Three other studies assessed the performance of the MoCA against other tools. Nantes et al (28) compared the performance of the MoCA against the HVLT-R in 98 patients with SLE. The MoCA showed a moderate correlation with the HVLT-R ($r = 0.42$, $P < 0.0001$, AUC of 0.71), with a sensitivity of 73%, and poor specificity of only 63%. The 2 other studies compared

the MoCA with the ANAM. Adhikari et al (30) evaluated a small group of 44 SLE patients and reported a moderate correlation ($r = 0.57$, $P < 0.001$) between the ANAM and MoCA, demonstrating a sensitivity of 83% and specificity 73% of the MoCA compared to the ANAM. They used the total throughput score for ANAM interpretation with a relatively stringent definition for CI; a patient was defined as CI when the total throughput score was >2 SDs below the mean of the control population. The ANAM total throughput score has been shown previously to have a low performance against the ACR-NB, irrespective of the CI definition by the ANAM (40). Using the total throughput score as the sole score to interpret the ANAM is limited (15). Other studies found that using several scores generated by the ANAM for ANAM interpretation, such as the percentage of correct responses, mean reaction time, and coefficient of variation of reaction time, in addition to the total throughput score, was better for detection of CI than the total throughput score use alone (17,41). Additionally, Adhikari et al (30) included RA patients as controls. Since RA patients experience CI as well (42), choosing RA patients for controls may lead to underestimation of CI in SLE patients, selecting only patients with severe CI. Chalhoub and Luggen (31) evaluated 78 SLE patients and showed that the MoCA and ANAM scores were moderately correlated ($r = 0.51$, $P < 0.001$), with poor sensitivities and specificities, 52 to 63% and 59 to 68%, respectively (depending on the definition of CI used for each analysis).

The MoCA is a tool designed to distinguish adults with and without mild neurocognitive disorder (conceptualized as being a transitional or risk state between normal cognitive aging and dementia) or mild dementia in the general population (19,43). Therefore, it gives a fair amount of weight to cognitive domains that are considered important components of CI in the older adult general population, such as orientation and language (19). Among patients with SLE and CI, common deficits are found in the domains of learning and memory, visuospatial construction, simple attention and processing, and psychomotor speed. Domains of orientation and language are not frequently involved (28,44,45). Accordingly, we found that MoCA tasks of abstraction, language, and orientation could not differentiate between CI and non-CI patients and are most likely of limited utility in the screening of CI in SLE patients. Some domains affected in non-SLE CI conditions are not relevant to an SLE population and thus limit the utility of the MoCA in its present form and scoring system in patients with SLE. Further, given that the MoCA was developed as a brief screening test to identify clinical levels of CI in older adults, it does not necessarily have the variability in item difficulty required to identify more subtle, but still meaningful, impairments in persons across the lifespan.

A limitation of our study is that it evaluated only individuals with sufficient English ability for completion of the ACR-NB, which may reduce the generalizability of the findings. Another limitation is the possibility of practice effects, because the ACR-NB and ANAM were completed after the MoCA. However, despite

measuring the same cognitive functions, the tasks of the ACR-NB, ANAM, and MoCA are not similar, and therefore, practice effects are unlikely. Nevertheless, order effects are possible, with the potential for more fatigue on the ANAM tasks, since it was completed after the MoCA and ACR-NB. This possibility may have decreased the cognitive performance of the ANAM, resulting in a higher impairment level in patients and may have increased the cognitive performance of the MoCA

This study cannot support the MoCA as a high-stake screening tool for CI in SLE patients. The MoCA does not have evidence for concurrent criterion validity compared to the ACR-NB. The MoCA failed to show the sensitivity and specificity needed for this application when compared to the ACR-NB, the gold standard. Future work could include different scoring algorithms on the MoCA or the development of a lupus-specific screening tool for CI. In the interim, clinicians and researchers may wish to choose a more detailed assessment that is sensitive to the presentation of CI in lupus.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Touma had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Tayer-Shifman, Yuen, Green, Kakvan, Katz, Bingham, Diaz Martinez, Ruttan, Wither, Tartaglia, Su, Bonilla, Choi, Appenzeller, Barraclough, Beaton, Touma.

Analysis and interpretation of data. Tayer-Shifman, Yuen, Green, Kakvan, Katz, Bingham, Diaz Martinez, Ruttan, Wither, Tartaglia, Su, Bonilla, Choi, Appenzeller, Barraclough, Beaton, Touma.

REFERENCES

1. Kaul A, Gordon C, Crow MK, et al. Systemic lupus erythematosus. Nat Rev Dis Primers 2016;2:16039.

2. Tayer-Shifman OE, Bingham KS, Touma Z. Neuropsychiatric systemic lupus erythematosus in older adults: diagnosis and management. Drugs Aging 2022;39:129–42.

3. ACR Ad Hoc Committee on Neuropsychiatric Lupus Nomenclature. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum 1999; 42:599–608.

4. Hanly JG, Kozora E, Beyea SD, et al. Nervous system disease in systemic lupus erythematosus: current status and future directions. Arthritis Rheumatol 2019;71:33–42.

5. Al Rayes HA, Tani C, Kwan A, et al. What is the prevalence of cognitive impairment in lupus and which instruments are used to measure it? A systematic review and meta-analysis. Semin Arthritis Rheum 2018;48: 240–55.

6. Kozora E, Hanly JG, Lapteva L, et al. Cognitive dysfunction in systemic lupus erythematosus: past, present, and future. Arthritis Rheum 2008;58:3286–98.

7. Petri M, Naqibuddin M, Carson KA, et al. Cognitive function in a systemic lupus erythematosus inception cohort. J Rheumatol 2008;35: 1776–81.

8. Tam LS, Wong A, Mok VC, et al. The relationship between neuropsychiatric, clinical, and laboratory variables and quality of life of Chinese patients with systemic lupus erythematosus. J Rheumatol 2008;35: 1038–45.

9. Sweet JJ, Doninger NA, Zee PC, et al. Factors influencing cognitive function, sleep, and quality of life in individuals with systemic lupus erythematosus: a review of the literature. Clin Neuropsychol 2004;18: 132–47.

10. Calderon J, Flores P, Aguirre JM, et al. Impact of cognitive impairment, depression, disease activity, and disease damage on quality of life in women with systemic lupus erythematosus. Scand J Rheumatol 2017;46:273–80.

11. Appenzeller S, Cendes F, Costallat LT. Cognitive impairment and employment status in systemic lupus erythematosus: a prospective longitudinal study. Arthritis Rheum 2009;61:680–7.

12. Ginsburg KS, Wright EA, Larson MG, et al. A controlled study of the prevalence of cognitive dysfunction in randomly selected patients with systemic lupus erythematosus. Arthritis Rheum 1992;35:776–82.

13. Kim MY, Sen D, Drummond RR, et al. Cognitive dysfunction among people with systemic lupus erythematosus is associated with reduced participation in daily life. Lupus 2021;30:1100–7.

14. Kozora E, Ellison MC, West S. Reliability and validity of the proposed American College of Rheumatology neuropsychological battery for systemic lupus erythematosus. Arthritis Rheum 2004;51:810–8.

15. Hanly JG, Omisade A, Su L, et al. Assessment of cognitive function in systemic lupus erythematosus, rheumatoid arthritis, and multiple sclerosis by computerized neuropsychological tests. Arthritis Rheum 2010;62:1478–86.

16. Hanly JG, Su L, Omisade A, et al. Screening for cognitive impairment in systemic lupus erythematosus. J Rheumatol 2012;39:1371–7.

17. Tayer-Shifman OE, Green R, Beaton DE, et al. Validity evidence for the use of Automated Neuropsychological Assessment Metrics as a screening tool for cognitive impairment in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2020;72:1809–19.

18. Yuen K, Beaton D, Bingham K, et al. Validation of the automated neuropsychological assessment metrics for assessing cognitive impairment in systemic lupus erythematosus. Lupus 2022;31:45–54.

19. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–9.

20. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.

21. Mikdashi JA, Esdaile JM, Alarcon GS, et al. Proposed response criteria for neurocognitive impairment in systemic lupus erythematosus clinical trials. Lupus 2007;16:418–25.

22. Shapiro AM, Benedict RH, Schretlen D, et al. Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. Clin Neuropsychol 1999;13:348–58.

23. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143: 29–36.

24. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: a practical guide to their development and use. Oxford: Oxford University Press; 2015.

25. Perneger TV. How to use likelihood ratios to interpret evidence from randomized trials. J Clin Epidemiol 2021;136:235–42.

26. McGee S. Simplifying likelihood ratios. J Gen Intern Med 2002;17: 646–9.

27. Kwiecien R, Kopp-Schneider A, Blettner M. Concordance analysis: part 16 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011;108:515–21.

28. Nantes SG, Su J, Dhaliwal A, et al. Performance of screening tests for cognitive impairment in systemic lupus erythematosus. J Rheumatol 2017;44:1583–9.

29. Yuen K, Green R, Bingham K, et al. Metrics and definitions used in the assessment of cognitive impairment in systemic lupus erythematosus: a systematic review. Semin Arthritis Rheum 2021;51:819–30.

30. Adhikari T, Piatti A, Luggen M. Cognitive dysfunction in SLE: development of a screening tool. Lupus 2011;20:1142–6.

31. Chalhoub NE, Luggen ME. Screening for cognitive dysfunction in systemic lupus erythematosus: the Montreal Cognitive Assessment Questionnaire and the Informant Questionnaire on Cognitive Decline in the Elderly. Lupus 2019;28:51–8.

32. Paez-Venegas N, Jordan-Estrada B, Chavarria-Avila E, et al. The Montreal Cognitive Assessment Test: a useful tool in screening of cognitive impairment in patients with systemic lupus erythematosus. J Clin Rheumatol 2019;25:325–8.

33. Kanapathy A, Nik Jaafar NR, Shaharir SS, et al. Prevalence of cognitive impairment using the Montreal Cognitive Assessment questionnaire among patients with systemic lupus erythematosus: a cross-sectional study at two tertiary centres in Malaysia. Lupus 2019;28:854–61.

34. Maciel RO, Ferreira GA, Akemy B, et al. Executive dysfunction, obsessive-compulsive symptoms, and attention deficit and hyperactivity disorder in systemic lupus erythematosus: evidence for basal ganglia dysfunction? J Neurol Sci 2016;360:94–7.

35. Butt BA, Farman S, Khan SE, et al. Cognitive dysfunction in patients with systemic lupus erythematosus. Pak J Med Sci 2017;33:59–64.

36. Yue R, Gurung I, Long XX, et al. Prevalence, involved domains, and predictor of cognitive dysfunction in systemic lupus erythematosus. Lupus 2020;29:1743–51.

37. Raghunath S, Glikmann-Johnston Y, Morand E, et al. Evaluation of the Montreal Cognitive Assessment as a screening tool for cognitive dysfunction in SLE. Lupus Sci Med 2021;8:e000580.

38. Muschelli J. ROC and AUC with a binary predictor: a potentially misleading metric. J Classif 2020;37:696–708.

39. Papastefanakis E, Dimitraki G, Ktistaki G, et al. Screening for cognitive impairment in systemic lupus erythematosus: application of the Montreal Cognitive Assessment (MoCA) in a Greek patient sample. Lupus 2021;30:2237–47.

40. Touma Z, Beaton D, Tartaglia C, et al. Can the Automated Neuropsychological Assessment Metrics (ANAM) predict cognitive impairment compared to a comprehensive neuropsychological battery in patients with lupus? [abstract]. 2017. URL: https://acrabstracts.org/abstract/can-the-automated-neuropsychological-assessment-metrics-anam-predict-cognitive-impairment-compared-to-a-comprehensive-neuropsychological-battery-in-patients-with-lupus/.

41. Brunner HI, Klein-Gitelman MS, Zelko F, et al. Validation of the Pediatric Automated Neuropsychological Assessment Metrics in childhood-onset systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2013;65:372–81.

42. Meade T, Manolios N, Cumming SR, et al. Cognitive impairment in rheumatoid arthritis: a systematic review. Arthritis Care Res (Hoboken) 2018;70:39–52.

43. Petersen RC, Caracciolo B, Brayne C, et al. Mild cognitive impairment: a concept in evolution. J Intern Med 2014;275:214–28.

44. Moghaddam B, Beaton D, Green R, et al. Prevalence of cognitive impairment in an inception lupus cohort as assessed by a comprehensive neuropsychological battery [abstract]. 2019. URL: https://acrabstracts.org/abstract/prevalence-of-cognitive-impairment-in-an-inception-lupus-cohort-as-assessed-by-a-comprehensive-neuropsychological-battery/.

45. Bertsias GK, Ioannidis JP, Aringer M, et al. EULAR recommendations for the management of systemic lupus erythematosus with neuropsychiatric manifestations: report of a task force of the EULAR standing committee for clinical affairs. Ann Rheum Dis 2010;69:2074–82.