ACR Open Rheumatology
Vol. 5, No. 9, September 2023, pp 499–507
DOI 10.1002/acr2.11588

American College of Rheumatology
Empowering Rheumatology Professionals

Remote Administration of Physical and Cognitive Performance Assessments in a Predominantly Black Cohort of Persons With Systemic Lupus Erythematosus

Courtney Hoge,[1] C. Barrett Bowling,[2,3] Charmayne Dunlop-Thomas,[1] Brad D. Pearce,[4] Cristina Drenkard,[1,4] S. Sam Lim,[1,4] and Laura C. Plantinga[5]

Objective. In a study of physical and cognitive functioning among predominantly Black individuals with systemic lupus erythematosus (SLE), we compared remotely administered physical and cognitive performance assessments to those collected in person.

Methods. A subset of participants who completed an in-person visit in our parent study from 2021 to 2022 (n = 30) were recruited to complete a second, remote visit within 28 days. Physical performance (measured by a modified Short Physical Performance Battery [SPPB]; range 0-12; subscale ranges 0-4; higher = better performance) and cognitive performance (episodic and working memory adjusted *t*-scores, measured using NIH Toolbox) were measured at both visits. Mean scores were compared using paired *t*-tests; intraclass correlation coefficients (ICCs) were obtained from two-way mixed effects models. Linear and logistic models were used to estimate stratified associations between performance measures and related outcomes.

Results. Participants were primarily female (93.3%) and Black (93.3%). In-person versus remote overall SPPB (8.76 vs. 9.43) and chair stand (1.43 vs. 1.90) scores were statistically significantly lower. *t*-Scores for episodic memory (47.27 vs. 49.53) and working memory (45.37 vs. 47.90) were lower for in-person versus remote visits. The ICC for overall SPPB indicated good agreement (0.76), whereas the ICCs for episodic (0.49) and working memory (0.57) indicated poor-moderate agreement. Associations between assessments of performance with related outcomes were similar and did not statistically significantly differ by modality of visit.

Conclusion. To possibly expand and diversify pools of participants in studies of physical and cognitive performance in SLE, remote administration of assessments should be considered for future research.

INTRODUCTION

The COVID-19 pandemic disrupted in-person human subjects research but also presented an opportunity for researchers to develop, administer, and compare remote versions of assessments previously performed in person. Remotely administered assessments of physical and cognitive performance would allow investigators to include individuals whose functioning and/or social circumstances might preclude traveling to and completing a study visit in future studies. Efforts to increase diversity in clinical trials particularly (1) are a national imperative. Despite the potential advantages of using such remote measures, the feasibility and comparability of remote versus in-person assessments of physical and cognitive performance remain unknown.

In this study (Approaches to Positive, Patient-centered Experiences of Aging in Lupus [APPEAL]), we assessed physical and

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Supported by the National Institute on Aging of the NIH under grant R01-AG-061179 (to Dr. Plantinga). The Georgians Organized Against Lupus (GOAL) cohort was supported by Centers for Disease Control and Prevention (CDC) grant 1U01-DP-006488 (to Drs. Lim and Drenkard) at the time of this study. The GOAL cohort is currently supported by CDC grant 1U01-DP-006698-01.

[1]Courtney Hoge, MSPH, Charmayne Dunlop-Thomas, MS, MPH: School of Medicine, Emory University, Atlanta, Georgia; [2,3]C. Barrett Bowling, MD, MSPH: Geriatric Research Education and Clinical Center, Durham Veterans Affairs, Durham, North Carolina and School of Medicine, Duke University, Durham, North Carolina; [4]Brad D. Pearce, PhD: Rollins School of Public Health, Emory University, Atlanta, Georgia; [1,4]Cristina Drenkard, MD, PhD, S. Sam Lim, MD, MPH: School of Medicine, Emory University, Atlanta, Georgia and Rollins School of Public Health, Emory University, Atlanta, Georgia; [5]Laura C. Plantinga, PhD: School of Medicine, University of California San Francisco, San Francisco, California.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr2.11588.

Address correspondence via email to Courtney Hoge, MSPH, at courtney.hoge@emory.edu.

Submitted for publication April 7, 2023; accepted in revised form June 28, 2023.

cognitive performance in both in-person and remote visits in a cohort of primarily Black individuals with systemic lupus erythematosus (SLE). Particularly for this population, participants often experience health-related demands in addition to life-related demands (eg, job responsibilities, childcare) that affect their ability to participate in in-person research (2,3). Having a feasible and comparable set of remote measures of functioning may result in opportunities for remote participation in research, which may result in fewer refusals to participate, and thus reduced selection bias and improved equity. Here, we sought to describe and compare physical and cognitive performance measures assessed remotely versus in person among a subset of participants who presented for an in-person APPEAL study visit and a subsequent remote study visit.

PATIENTS AND METHODS

Study population. For this comparative substudy, 30 APPEAL participants who completed in-person study visits between November 18, 2021, and March 3, 2022, were recruited via convenience sampling to complete a second, remote visit (number of days between visits: 4 to 28 days), conducted via Zoom (San Jose, CA). In this substudy, the remote visit included physical performance and cognitive functioning assessments that had been modified for remote administration in the overall study. APPEAL participants were recruited from the ongoing, population-based Georgians Organized Against Lupus (GOAL) cohort of adults (≥18 years) living in metropolitan Atlanta with a documented diagnosis of SLE, details of which have been published previously (4–6). Inclusion criteria for the APPEAL ancillary study (N = 451; 206 in-person study visits and 245 remote study visits) were as follows: the ability to speak English, sufficient vision and hearing to undergo study testing, ability to consent, and living in Georgia at the time of the study visit. For the substudy reported here, lack of access to a smartphone or computer (preferred) with internet access and a camera was an additional exclusion criterion. All participants provided informed consent for both study visits. The Emory Institutional Review Board and Grady Research Oversight Committee approved the APPEAL study protocol.

Development and administration of remote functioning measures. Due to COVID-19-related safety concerns, recruitment for in-person APPEAL visits, which had begun in October 2019, was halted in March 2020. During the recruitment pause, the study team converted study assessments that previously required in-person administration (and were not collected by survey) to remotely administered assessments. Each element of the study was thoroughly tested within the study team, as part of complete mock study visits via Zoom, and piloted among GOAL participants (n = 7) who were not eligible for APPEAL due to current residence outside of Georgia. This process allowed us to identify additional issues that might arise with real study participants, further train study personnel on new methods, and finalize remote study procedures and materials. This modified study protocol was used for remote visits in the comparative substudy reported here.

Prior to and during remote visits, participants were instructed to wear shoes, identify a chair for chair stands (preferably with a hard surface and a back, and no arms; wheels were not allowed), and locate a space large enough to stand and walk in place. The balance and chair stand tasks were performed as in the in-person Short Physical Performance Battery (SPPB), with the study coordinator timing the tasks remotely. Given the limitations of participants' living spaces and inability to measure a standard walking course, participants completed a 1-minute walk-in-place in place of a timed 4 m walk. Study coordinators counted the number of steps completed with hand counters. Steps were counted if participants' knees reached a 45-degree angle (ie, foot entirely cleared the ground); participants were provided with a sticky note, which was mailed to participants for remote visits and given to participants after their in-person visit for remote validation visits, to "mark" this position on the wall behind them prior to beginning the task. Brief videos illustrating each of the three tasks (including foot positions for balance tasks, wall marking for the walk-in-place, and proper arm position for chair stands) were shown to the participant just prior to the task. Information on potential mitigating environmental factors (eg, chairs and footwear) was recorded by study coordinators as well. A comparison of the administration of the SPPB in remote versus in-person visits can be found in Table 1.

Remote NIH Toolbox episodic and working memory tasks were administered via an iPad, as during in-person visits. The video and audio from the iPad were broadcast to the Zoom meeting during testing by the coordinator, who manipulated the iPad as needed based on the participants' instructions. A comparison of the administration of the NIH Toolbox measures in remote versus in-person visits can be found in Table 2.

Variables. *Physical performance: SBBP.* Physical performance was assessed via the SPPB (7), which includes assessments of balance, gait speed, and lower body strength.

All individual assessments for both in-person and remote visits were scored 0 to 4 (with higher scores indicating higher levels of physical performance), based on the cutoffs from the SPPB (7); the overall score was the sum of the three individual scores (range, 0-12). For remote visit gait speed, walk-in-place steps were converted to distance walked using the conversions 0.413*(height in meters) for female participants and 0.415*(height in meters) for male participants (8). Gait speed for the remote visits was then estimated as (estimated distance in meters)/(60 seconds).

Cognitive performance: episodic and working memory. The two cognition assessments that could be administered remotely (9), via the NIH Toolbox application (10–12), were the Picture Sequence Memory Test (measures episodic memory, or ability to remember objects, people, or events experienced at

Table 1. Summary of administration of the SPPB for in-person versus remote study visits

In-person administration	Remote administration
Initial setup	
• Mark 4-m course on floor (left in place for all visits). • Gather timer, study documents for recording values, and armless folding chair (same chair used for all visits).	• Gather timer, handheld tally counter, study documents for recording values. • Work with the participant to properly position the camera to be able to see the participant perform tasks. • Ensure the participant is wearing appropriate footwear and has a chair without wheels and documents.
Balance	
• Provide verbal instructions for tasks. • Demonstrate tasks for participant. • Record time in seconds for remaining in stand position, without stepping out or using arms to steady (truncated at 10 seconds). • Document reasons for incomplete testing as applicable.	• Provide verbal instructions for tasks. • Share videos of demonstrations (including close-ups of foot positions for side-by-side, semitandem, and tandem stands). • Record time in seconds for remaining in stand position, without stepping out or using arms to steady (truncated at 10 seconds). • Additionally note any issues that may result in invalid data (eg, problems with internet connection during testing).
Gait speed	
• Provide verbal instructions for task. • Demonstrate task for participant. • Record time in seconds for two 4-m course walks performed at usual speed. • For scoring, calculate gait speed for the fastest of the two walks. • Document reasons for incomplete testing as applicable.	• Provide verbal instructions for task. • Share video demonstration of task (including placement of sticky note on wall to indicate knee position with foot entirely clearing the ground). • Counts steps (using handheld tally counter) for 1-minute walk-in-place at usual speed. • Additionally note any issues that may result in invalid data (eg, problems with internet connection during testing). • For scoring, estimate gait speed estimated as (number of steps*estimated step length[a])/60 seconds.
Chair stands	
• Provide verbal instructions for task. • Demonstrate task for participant. • Assess whether participant can complete a single chair stand. • If so, record time in seconds to completed five chair stands. • Document reasons for incomplete testing as applicable.	• Provide verbal instructions for tasks. • Share video demonstration of task (including crossing of arms over chest and sitting all the way back in chair). • Assess whether participant can complete a single chair stand. • If so, record time in seconds to completed five chair stands. • Additionally note any issues that may result in invalid data (eg, problems with internet connection during testing).

Abbreviations: SPPB, Short Physical Performance Battery.
[a]Step length estimated as 0.413*(height in m) for female participants and 0.415*(height in m) for male participants.

particular times and places) and the List Sorting Working Memory Test (measures working memory, or the ability to remember and see connections between items or ideas), neither of which require measurement of response time. All individual raw Toolbox scores were converted to *t*-scores, which were adjusted for age, sex, race, ethnicity, and education. Fully adjusted *t*-scores (mean = 50, SD = 10) range from 0 to 100, such that 50 is the average score and 40 and 60 are 1 SD below and above the mean, respectively; higher scores indicate better cognitive functioning.

Self-reported health. Participants self-reported their relative health on the day of in-person and remote assessments via a single item ("Compared to your usual health over the past year, how do you feel today?" with potential responses of much worse than usual [=1], worse than usual [=2], about the same [=3], better than usual [=4], and much better than usual [=5]).

Related outcomes. For physical performance, the outcomes of falls and self-reported physical function were examined. Participants self-reported whether they had fallen in the last 12 months (yes or no). Self-reported physical functioning was obtained via the Patient Reported Outcomes Measurement Information System (PROMIS) Physical Functioning-Short Form 12a (13) and converted to *t*-scores. For cognitive performance, self-reported forgetfulness and Trail Making Test times (a measure of executive function) were examined as outcomes. Self-reported forgetfulness was obtained from a single item on the Systemic Lupus Activity Questionnaire (SLAQ) item (mild, moderate, or severe vs. no problem in the last 3 months), and Trail Making B (14) times were obtained in person and truncated at 5 minutes (300 seconds).

Other variables. All other variables of interest were assessed during the initial in-person study assessment. Participant demographics (age, gender, race, ethnicity, and education) were obtained via NIH Toolbox. SLE activity was assessed via the self-administered SLAQ (range 0 to 44; higher scores indicating greater SLE-related disease activity) (15). Height and weight were measured, and body mass index (BMI) was calculated as weight in kg divided by height in meters squared. Depression was assessed via the self-administered PROMIS Depression Short Form-8a (16) on the day of the in-person study visit. Perceived stress was assessed via the self-administered Cohen's 10-item Perceived Stress Scale (range, 0 to 40; higher scores representing greater perceived stress) (17,18). Participants self-reported whether they were currently taking steroids (yes or no).

Table 2. Summary of administration of the NIH Toolbox[a] Picture Sequence[b] and List Sorting[c] tasks for in-person versus remote study visits

In-person administration	Remote administration
Initial setup	
• Enter required demographic information for NIH Toolbox cognitive testing.	• Join the video call with the iPad and share screen and audio. • Enter required demographic information for NIH Toolbox cognitive testing.
Picture Sequence Memory Test	
• Initiate the Picture Sequence module on the iPad. • Allow participant to use the iPad to move pictures to complete the module tasks.	• Initiate the Picture Sequence module on the iPad. • Move pictures on the iPad per the participant's instructions to complete the module tasks. • Note any issues that may result in invalid data (eg, participant appears to be taking notes; another person in room providing help; problems with internet connection during testing).
List Sorting Test	
• Initiate the List Sorting module on the iPad. • Use the keyboard associated with the iPad to indicate correctness of verbal participant responses.	• Initiate the List Sorting module on the iPad. • Use the keyboard associated with the iPad to indicate correctness of verbal participant responses. • Note any issues that may result in invalid data (eg, participant appears to be taking notes; another person in room providing help; problems with internet connection during testing).

Abbreviation: NIH, National Institutes of Health.
[a]Set of standardized measures assessing cognitive, emotional, motor, and sensory function.
[b]Measures episodic memory: ability to remember objects, people, or events experienced at particular times and places.
[c]Measures working memory: ability to remember and see connections between items or ideas.

Statistical analysis. Participant characteristics were summarized overall and by type of visit using *t*-test or χ^2 test, as appropriate. For each of the assessments that were performed both in person and remotely, mean values were compared by paired *t*-test. Intraclass correlation coefficients (ICCs) to assess test-retest reliability were estimated using two-way mixed effects models with absolute agreement. Betas and odds ratios from crude linear and logistic regression models stratified by visit type were used to assess whether assessments of physical and cognitive performance were associated similarly with related outcomes; an interaction term between the performance measures and mode of administration (remote vs. in person) was used to determine whether associations differed by type of study visit. Analyses were performed using Stata v. 17 (College Station, TX).

RESULTS

Characteristics of substudy participants. Similar to the overall APPEAL sample, participants who completed both an in-person and a remote study visit (ie, were included in the substudy) had a mean (SD) age of 46.2 (11.9) years, were predominately female (93.3%), self-identified as being Black (93.3%), and were not currently employed (63.0%) (Table 3). Participants who were not in the substudy sample had a mean (SD) age of 45.7 (11.4) years and were predominately female (91.7%), Black (81.7%), and not currently employed (52.2%). Participants in the substudy largely had some college or associates degree (43.3%) as their highest level of education. The mean (SD) SLAQ score for substudy versus non-substudy participants was 11.9 (6.6) vs. 11.8 (7.6). The mean (SD) BMI for participants in the validation substudy was slightly higher than for participants who were not in the substudy (32.6 [9.2] kg/m^2 vs. 30.0 [8.1] kg/m^2). Mean (SD) depression *t*-scores (50.6 [9.6] vs. 47.9 [9.1]) and perceived stress scores (15.7 [7.8] vs. 15.2 [7.1]) were similar for those who were in the substudy versus those who were not in the substudy. The only characteristic that statistically significantly differed by whether participants were included in the substudy was reported current use of steroids: 23.3% of participants in the substudy reported currently taking steroids versus 44.1% of participants not in the substudy (Table 3).

Comparison of assessments between in-person and remote study visits. The mean (SD) days between the in-person study visit and remote study visit was 17.4 (7.0). Participants answered the current health status question for in-person and remote study visits similarly (mean [SD] scores: 3.20 [0.61] and 3.20 [0.76], $P = 0.88$) (Table 4): 15 participants reported the same level of health at both visits (days between visits: 4-23), 12 participants had a 1-level difference in their health between visits (days between visits: 6-28), and 3 reported a 2-level difference in their health between visits (days between visits: 9-17).

Physical performance. Overall SPPB scores statistically significantly differed between the two visits, with scores for in-person study visits having a mean (SD) of 8.76 (2.08) and scores for remote study visits having a higher mean (SD) of 9.43 (1.94) (Table 4). Of the three SPPB domains, only chair stand scores statistically significantly differed between visits, with in-person visit scores being lower than remote study visit scores (mean [SD]: 1.43 [0.97] vs. 1.90 [1.35]). The overall SPPB scores had good agreement (ICC: 0.76, 95% CI: 0.50-0.88) with the balance subscore having excellent agreement (ICC: 0.91, 95% CI: 0.82-0.96); however, both gait speed score and actual versus estimated gait speed had poor agreement (Table 4).

Cognitive performance. The mean (SD) episodic memory adjusted *t*-scores for in-person and remote study visits were 47.27 (8.14) and 49.53 (9.26), respectively, with poor agreement (ICC: 0.49, 95% CI: 0.17-0.72) (Table 4). After removing data for

Table 3. Comparison of characteristics of participants overall and by participation in current substudy

Characteristic	APPEAL study participants			
	Overall	In substudy	Not in substudy	P[a]
N	451	30	421	
Dates of visits and surveys	10/8/19 to 5/12/22	11/18/21 to 3/3/2022	10/8/19 to 5/12/22	
Sociodemographics				
Mean (SD) age, y	46.2 (11.8)	46.2 (11.9)	45.7 (11.4)	0.8
Age category, n (%)				0.9
18-34	92 (20.4%)	6 (20.0%)	86 (20.4%)	
35-49	174 (38.6%)	11 (36.7%)	163 (38.7%)	
≥50	185 (41.0%)	13 (43.3%)	172 (40.9%)	
Gender,[b] n (%)				0.8
Female	414 (91.8%)	28 (93.3%)	386 (91.7%)	
Male	37 (8.2%)	2 (6.7%)	35 (8.3%)	
Race, n (%)				0.3
Black	372 (82.5%)	28 (93.3%)	344 (81.7%)	
White	52 (11.5%)	1 (3.3%)	51 (12.1%)	
Other	27 (6.0%)	1 (3.3%)	26 (6.2%)	
Ethnicity, n (%)				0.6
Hispanic	25 (5.6%)	1 (3.3%)	24 (5.7%)	
Not Hispanic	425 (94.4%)	29 (96.7%)	396 (94.3%)	
Highest level of education completed, n (%)				0.4
<High school	20 (4.4%)	1 (3.3%)	19 (4.5%)	
High school degree or equivalency	83 (18.4%)	8 (26.7%)	75 (17.8%)	
Some college or associates degree	173 (38.4%)	13 (43.3%)	160 (38.0%)	
College graduate	102 (22.6%)	3 (10.0%)	99 (23.5%)	
Postcollege	73 (16.2%)	5 (16.7%)	68 (16.2%)	
Currently employed, n (%)				0.3
Yes	207 (47.2%)	10 (37.0%)	197 (47.8%)	
No	232 (52.9%)	17 (63.0%)	215 (52.2%)	
Clinical				
Mean (SD) SLAQ score	11.8 (7.5)	11.9 (6.6)	11.8 (7.6)	0.9
Mean (SD) BMI, kg/m^2	30.1 (8.2)	32.6 (9.2)	30.0 (8.1)	0.09
Mean (SD) depression t-score[c]	48.0 (9.2)	50.6 (9.6)	47.9 (9.1)	0.1
Mean (SD) perceived stress score[d]	15.2 (7.1)	15.7 (7.8)	15.2 (7.1)	0.7
Currently taking steroids, n (%)	192 (42.7%)	7 (23.3%)	185 (44.1%)	0.03

Abbreviations: APPEAL, Approaches to Positive, Patient-Centered Experiences of Aging in Lupus; BMI, body mass index; GOAL, Georgians Organized Against Lupus (parent study); PROMIS, Patient Reported Outcomes Measurement Information System; SLAQ, Systemic Lupus Activity Questionnaire (range, 0-47; 47 is maximum activity).
[a]For in-person versus remote APPEAL participants, by t-test or χ^2 test, as appropriate.
[b]Represents sex at birth for APPEAL participants.
[c]From the PROMIS Depression Short Form-8a.
[d]From Cohen's 10-item Perceived Stress Scale (range, 0-40; higher scores representing greater perceived stress).

two participants with potentially invalid cognition scores (see Table 2) due to suspected notetaking and internet interruptions, episodic memory scores were fairly similar to scores including all validation study participants (in-person and remote mean [SD] t-scores: 47.18 [8.43] and 48.14 [7.57], respectively), and there was moderate agreement (ICC: 0.61, 95% CI: 0.31-0.80). To address possible issues with test administration, episodic memory t-scores that changed by more than 1 SD (n = 6) were removed, which resulted in nearly identical scores for in-person and remote study visits (mean [SD]: 48.08 [8.71] and 48.08 [7.26]) and good agreement (ICC: 0.78, 95% CI: 0.56-0.90) (Table 4).

Working memory t-scores for in-person and remote study visits were 45.37 (10.76) and 47.90 (9.15), respectively (Table 4). Working memory scores had moderate agreement (ICC: 0.57; 95% CI: 0.28-0.77) and did not statistically significantly differ between in-person and remote study visits. After removing data for one participant with a potentially invalid score (see Table 2) due to suspected notetaking, mean (SD) working memory t-scores were 45.00 (10.76) for in-person study visits and 46.86 (7.30) for remote study visits; the ICC did not substantially change (ICC: 0.57; 95% CI: 0.27-0.77). Removing participants' working memory t-scores because of more than 1 SD change in scores (n = 7) between study visits gave similar results (in-person vs. remote, mean [SD]: 45.13 [8.15] vs. 46.17 [7.97]); however, there was good agreement for working memory scores (ICC: 0.80; 95% CI: 0.59-0.91) between visits when these data were removed (Table 4).

Table 4. Comparison of measurements between in-person and remote study visits

Measure	Mean (SD)		P[a]	ICC (95% CI)[b]
	In person	Remote		
Current health status[c]	3.20 (0.61)	3.20 (0.76)	0.9	—
SPPB[d] overall score	8.76 (2.08)	9.43 (1.94)	0.01	0.76 (0.50 to 0.88)
Balance score	3.80 (0.76)	3.83 (0.75)	0.6	0.91 (0.82 to 0.96)
Gait speed score	3.53 (0.90)	3.70 (0.60)	0.3	0.23 (−0.14 to 0.54)
Actual and estimated gait speed[e]	0.96 (0.20)	0.93 (0.27)	0.5	0.48 (0.14 to 0.72)
Chair stands	1.43 (0.97)	1.90 (1.35)	0.01	0.66 (0.36 to 0.83)
Episodic memory[f]	47.27 (8.14)	49.53 (9.26)	0.2	0.49 (0.17 to 0.72)
Working memory[f]	45.37 (10.76)	47.90 (9.15)	0.1	0.57 (0.28 to 0.77)

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient.
[a]By χ^2 or *t*-test, as appropriate.
[b]Using two-way mixed effects model with absolute agreement.
[c]Participants responded to the question "compared to your usual health over the past year how would you say you feel today?" on a Likert scale (1 = much worse than usual, 5 = much better than usual).
[d]SPPB, Short Physical Performance Battery.
[e]Remote study visit gait speed was estimated using gender- and height-specific estimates.
[f]From the NIH Toolbox Picture Sequence Memory and List Sorting Tests.

Table 5. Association of physical and cognitive performance measures with related outcomes, by administration type, among participants who completed both in-person and remote study visits

Exposure	Related outcome	Effect estimate (95% CI) for association		P[a]
		In person	Remote	
Physical performance measures				
SPPB overall score	Self-reported physical functioning[b]	2.34 (0.72 to 3.97)	2.50 (0.75 to 4.24)	0.9
	Falls[c]	0.73 (0.44 to 1.19)	0.84 (0.57 to 1.24)	0.7
Balance score	Self-reported physical functioning[b]	1.90 (−3.13 to 6.94)	1.88 (−3.26 to 7.01)	0.9
	Falls[c]	0.44 (0.07 to 2.61)	—[d]	0.9
Gait speed score	Self-reported physical functioning[b]	2.82 (−1.35 to 6.98)	4.98 (−1.23 to 11.18)	0.6
	Falls[c]	0.43 (0.14 to 1.34)	1.45 (0.39 to 5.34)	0.2
Gait speed, m/s	Self-reported physical functioning[b]	19.47 (0.74 to 38.20)	9.49 (−5.39 to 24.37)	0.4
	Falls[c]	0.04 (0.00 to 2.88)	3.07 (0.16 to 60.02)	0.1
Chair stand score	Self-reported physical functioning[b]	7.15 (4.28 to 10.02)	3.63 (1.13 to 6.14)	0.07
	Falls[c]	0.78 (0.36 to 1.69)	0.81 (0.46 to 1.41)	0.9
Cognitive performance measures				
Episodic memory	Self-reported forgetfulness[e]	1.01 (0.91 to 1.11)	1.00 (0.92 to 1.08)	0.9
	Trail Making B time[f]	−1.83 (−3.66 to −0.00)	−1.41 (−3.05 to 0.22)	0.7
Working memory	Self-reported forgetfulness[e]	0.96 (0.89 to 1.04)	0.93 (0.84 to 1.03)	0.6
	Trail Making B time[f]	−1.85 (−3.15 to −0.55)	−1.32 (−2.99 to 0.35)	0.6

Abbreviations: CI, confidence interval; OR, odds ratio; PROMIS, Patient Reported Outcomes Measurement Information System; SLAQ, Systematic Lupus Activity Questionnaire; SPPB, Short Physical Performance Battery.
[a]For interaction of exposure × modality of visit (remote vs. in person).
[b]*t*-Scores from the PROMIS Physical Functioning-Short Form 12a. Effect estimates represent betas from crude linear regression models.
[c]Reported falls (yes or no) in the last 12 months. Effect estimates represent ORs from crude logistic regression models.
[d]Not estimable.
[e]Forgetfulness from the SLAQ item (mild, moderate, or severe vs. no problem in the last 3 months). Effect estimates represent ORs from crude logistic regression models.
[f]Trail Making B times, in seconds. Effect estimates represent betas from crude linear regression model.

Association of performance measures with related outcomes. *Physical performance.* A 1-point increase in the overall higher SPPB scores were associated with 2.3- and 2.5-point higher *t*-scores (per +1 point on the SPPB) for self-reported physical functioning for in-person (β: 2.34, 95% CI: 0.72-3.97) and remote (β: 2.50, 95% CI: 0.75-4.24) study visits (Table 5). Higher balance scores were associated with higher self-reported physical functioning for in-person and remote study visits (β 1.90 [95% CI: –3.13 to 6.94] and β 1.88 [95% CI: –3.26 to 7.01], respectively). Both higher gait speed scores and gait speed in m/s were associated with higher self-reported physical functioning; however, the effect estimate (β [95% CI]) was stronger for gait speed (in-person vs. remote: β: 19.47 [95% CI: 0.74-38.20] vs. β: 9.49 [95% CI: –5.39 to 24.37] per +1 m/s faster gait speed). Higher chair stand score (+1 point) was associated with 7.15 (95% CI: 4.28-10.02) and 3.63 (95% CI: 1.13-6.14) higher self-reported physical functioning *t*-scores for in-person and remote study visits, respectively. None of the associations between SPPB scores, overall and by domain, and self-reported physical functioning were statistically significantly different by type of study visit (Table 5).

Higher SPPB scores, overall and by domain, were associated with a 27% and 16% lower risk of falls for in-person and remote study visits, respectively; however, none of the effect estimates were statistically significant, and the associations did not statistically significantly differ by type of study visit (Table 5). The OR (95% CI) for the association between chair stand score and falls in the last year for in-person and remote study visits was 0.78 (0.36-1.69) and 0.81 (0.46-1.41), respectively. For gait speed score and gait speed (OR [95% CI]), respectively, higher scores were associated with a lower risk of falls for in-person study visits (0.43 [0.14-1.34] and 0.04 [0.00-2.88]) but a higher risk of falls in remote study visits (1.45 [0.39-5.34] and 3.07 [0.16-60.02]) (Table 5).

Cognitive performance. t-Scores for episodic and working memory were not associated with self-reported forgetfulness for either type of study visit (Table 5). Associations of both episodic memory and working memory were negatively associated with Trail Making Test times, with each unit of *t*-score being associated with ∼1.4 to 1.8 seconds faster Trail Making Test times; the results were similar across study types and were not statistically significant (Table 5). None of the associations between cognitive performance and either forgetfulness or Trail Making Test times statistically significantly differed by modality of study visit (Table 5).

DISCUSSION

Overall, we found that remotely administered assessments of physical and cognitive performance were feasible and comparable in our study population of predominantly Black individuals with SLE. We found similar results for remotely administered versus in-person assessments of physical and cognitive performance: the overall SPPB score had good agreement, and episodic memory and working memory had poor to moderate agreement. Importantly, associations of both physical and cognitive scores with related outcomes were generally similar across study types, suggesting that, although scores may differ across study types, their relative predictive ability may stay fairly consistent. Although our study was a convenience study and a formal validation study is still needed, our findings provide initial evidence that remote assessments may be comparable to in-person assessments in an SLE population, which is generally younger and subject to fluctuations in disease status, compared with other populations among whom such validation studies have been performed previously (19–21).

We found that all SPPB subscores indicated better performance in remote visits; however, only chair stand scores were statistically significantly higher. Additionally, SPPB overall and all subscores were positively associated with self-reported physical functioning for both visit types. Similar to other studies (20), balance scores did not significantly differ between in-person and remote study visits and had excellent agreement. Although gait speed scores had the lowest agreement of the SPPB measures, these scores, and actual versus estimated gait speed, did not statistically significantly differ between visits. Chair stand scores were statistically significantly higher for remote visits in comparison to in-person visits, which could be attributed to the use of a standardized chair for in-person visits. In remote study visits, participants used chairs that were available to them, which may have had cushioning or been an inappropriate height, making completing chair stands easier. Although task learning and conserved energy not spent on traveling to the study site may have also contributed to the higher scores in the remote visits overall, they are unlikely to explain the more substantial difference in chair stand scores. Thus, documentation of the type of chair may be important for remote administration of the SPPB, because it could allow for sensitivity analyses excluding data points that might not be comparable.

Both episodic and working memory scores were similar between in-person and remote study visits in our substudy, but with poor and moderate agreement, respectively, which is similar to another study examining the validity of administering the NIH Toolbox remotely (22). Given the lack of standardized environments in remote visits, we performed sensitivity analyses removing participants who may have had invalid scores due to internet interruptions or possible notetaking, which we documented. These results showed greater agreement, suggesting an observant coordinator and careful exclusions could increase the reliability of remotely administered cognitive assessments. We also removed scores that changed by more than 1 SD between visits, which would not be possible in a real-time study with single-mode visits. Although this exclusion improved the reliability of the measures, it did not substantially change the mean scores, suggesting that, on average, scores will be similar with remote administration.

It is important to note that, to our knowledge, remote assessment of physical and cognitive function, not including patient reported outcomes related to functioning (23), has not yet been reported specifically in SLE populations. Prior studies in cancer survivors (19,20) and healthy older adults (21) have taken similar but not identical approaches to remote measurement of physical performance. For example, none of the studies used a walk-in-place substitution for walking speed and instead had participants walk measured courses (19,20); however, this required having inclusion criteria of having at least 12 feet of space and the ability to clear obstacles (20), which we did not feel comfortable imposing given the socioeconomic limitations of many GOAL participants. Timed up-and-go tests were performed in one study (21), which we did not perform, also due to potential space limitations. Additionally, in one study (20), dyads of patients and caregivers were used, such that each participant had someone to assist in setting up courses and webcams; while some of our participants had others in the home to provide this type of help, we found that these helpers could sometimes intervene during cognitive testing. Requiring another person's presence would have further limited the generalizability of our study. These prior studies also had shorter time periods between remote and in-person studies (within 1 week), which we were not able to accomplish with our limited study resources and among our population, who often had work and family responsibilities that might have been less frequent among older cancer survivors (19,20) and healthy older adults (21). These individuals were probably also less likely than individuals with SLE to have sudden fluctuations in signs and symptoms that might affect functioning. Finally, similarly to our study, prior studies (19–21) did not use specific apps to measure time and distance for physical performance testing. However, studies including such apps, which might eliminate the need for providing materials and lower costs to and burden on the participant, could be considered for future studies.

There are several additional limitations that deserve mentioning. First, this substudy was completed to address concerns of using modified, remote measures and was not an original aim of the study; thus, no conclusions about the validity of measures can be made. The small sample size precludes adjustment for confounding factors and is subject to greater random error. Participants who completed remote study visits required access to a laptop or smartphone with a camera, the internet, and Zoom, which possibly leads to selection bias. With remote administration of measures, the environment cannot be controlled; for example, internet interruptions and delays may affect physical performance measures, and participants may take notes during cognitive assessments. Additionally, home environments may contain barriers or distractions that are not present in our physical study space. Although all remote visits in this validation study were performed by the same coordinator (C.H.; with another, single coordinator performing the in-person counterpart visits), both in-person and remote study visits for our parent study were performed by multiple coordinators, increasing potential variability in measurement within and across administration type.

In conclusion, data for some physical and cognitive performance measures can feasibly be collected remotely in patients with SLE. Such data could be reasonably combined with data collected in person, with the addition of sensitivity analyses that examine within-mode results and/or exclude values that are potentially invalid for documented reasons. More broadly, the option of remote administration of physical and cognitive measures in research studies could be considered for future studies, including those with predominantly Black populations, similar to our study. These remote measures could diversify and increase the pool of potential participants by making studies more accessible to populations with historically greater barriers to in-person study participation, those with socioeconomic challenges, and those with impaired activities of daily living that limit in-person participation in studies, all of whom are overrepresented in the SLE population.

ACKNOWLEDGMENTS

We thank the participants of the APPEAL study. We also thank Sydnei Simpson and Olivia Barnum for collecting and validating data.

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 published. Hoge and Plantinga had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Hoge, Bowling, Pearce, Drenkard, Lim, Plantinga.

Acquisition of data. Hoge, Dunlop-Thomas, Drenkard, Lim, Plantinga.

Analysis and interpretation of data. Hoge, Plantinga.

REFERENCES

1. Kelsey MD, Patrick-Lake B, Abdulai R, et al. Inclusion and diversity in clinical trials: actionable steps to drive lasting change. Contemp Clin Trials 2022;116:106740.
2. Drenkard C, Easley K, Bao G, et al. Overcoming barriers to recruitment and retention of African-American women with SLE in behavioural interventions: lessons learnt from the WELL study. Lupus Sci Med 2020;7:e000391.
3. Falasinnu T, Bao G, Brady TJ, et al. Factors associated with the initiation and retention of patients with lupus in the chronic disease self-management program. Arthritis Care Res (Hoboken) 2023;75:519–28.
4. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725.
5. Lim SS, Bayakly AR, Helmick CG, et al. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: The Georgia Lupus Registry. Arthritis Rheumatol 2014;66:357–68.
6. Drenkard C, Rask KJ, Easley KA, et al. Primary preventive services in patients with systemic lupus erythematosus: study from a

population-based sample in Southeast U.S. Semin Arthritis Rheum 2013;43:209–16.

7. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49:M85–94.
8. Grieve DW, Gear RJ. The relationships between length of stride, step frequency, time of swing and speed of walking for children and adults. Ergonomics 1966;9:379–99.
9. NIH Toolbox. Remote Administration Guidelines for the NIH Toolbox: Response to COVID-19. 2020.
10. NIH Toolbox. Cognition [NIH Toolbox Cognition Batteries]. 2017. URL: http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox/intro-to-nih-toolbox/cognition.
11. Gershon RC, Cella D, Fox NA, et al. Assessment of neurological and behavioural function: the NIH Toolbox. Lancet Neurol 2010;9:138–9.
12. Weintraub S, Dikmen SS, Heaton RK, et al. The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: validation in an adult sample. J Int Neuropsychol Soc 2014;20:567–78.
13. Northwestern University. Intro to PROMIS. 2017. URL: https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis.
14. Corrigan JD, Hinkeldey NS. Relationships between parts A and B of the Trail Making Test. J Clin Psychol 1987;43:402–9.
15. Karlson EW, Daltroy LH, Rivest C, et al. Validation of a Systemic Lupus Activity Questionnaire (SLAQ) for population studies. Lupus 2003;12:280–6.
16. Hitchon CA, Zhang L, Peschken CA, et al. Validity and reliability of screening measures for depression and anxiety disorders in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72:1130–9.
17. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24:385–96.
18. Ezzati A, Jiang J, Katz MJ, et al. Validation of the perceived stress scale in a community sample of older adults. Int J Geriatr Psychiatry 2014;29:645–52.
19. Blair CK, Harding E, Herman C, et al. Remote assessment of functional mobility and strength in older cancer survivors: protocol for a validity and reliability study. JMIR Res Protoc 2020;9:e20834.
20. Hoenemeyer TW, Cole WW, Oster RA, et al. Test/retest reliability and validity of remote vs. in-person anthropometric and physical performance assessments in cancer survivors and supportive partners. Cancers (Basel) 2022;14:1075.
21. Peyrusqué E, Granet J, Pageaux B, et al. Assessing physical performance in older adults during isolation or lockdown periods: web-based video conferencing as a solution. J Nutr Health Aging 2022;26:52–6.
22. Rebchuk AD, Deptuck HM, O'Neill ZR, et al. Validation of a novel telehealth administration protocol for the NIH loolbox-cognition battery. Telemed J E Health 2019;25:237–42.
23. Bell K, Dykas C, Muckian B, et al. Patient-reported outcome information collected from lupus patients using a mobile application: compliance and validation. ACR Open Rheumatol 2022;4:99–109.