Published in final edited form as:

J Am Geriatr Soc. 2014 October ; 62(10): 1928–1932. doi:10.1111/jgs.13027.

Comparison of In-Person and Telephone Administration of the Mini-Mental State Examination in the University of Alabama at Birmingham Study of Aging

Richard E. Kennedy, MD, PhD^a, Courtney P. Williams, MPH^a, Patricia Sawyer, PhD^b, Richard M. Allman, MD^d, and Michael Crowe, PhD^c

^aDivision of Gerontology, Geriatrics, and Palliative Care, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama

^bComprehensive Center for Healthy Aging, Gerontology Education Program, University of Alabama at Birmingham, Birmingham, Alabama

^cDepartment of Psychology, University of Alabama at Birmingham, Birmingham, Alabama

^dGeriatrics and Extended Care Services, Veterans Health Administration, Washington, District of Columbia

Abstract

Objectives—To validate the Mini-Mental State Examination (MMSE) telephone (MMSET) against the MMSE.

Design—Cross-sectional.

Setting—Homes of community-dwelling older adults.

Participants—African-American and non-Hispanic white adults aged 75 and older participating in the University of Alabama at Birmingham Study of Aging II, a longitudinal epidemiological study across the state of Alabama (N=419).

Measurements—Cognition, measured using the MMSE, MMSET, and Six-Item Screener (SIS), and function, based on self-reported difficulty performing instrumental activities of daily living (IADLs). Correlation and agreement coefficients were used to examine concordance of the MMSE

Corresponding Author: Richard E. Kennedy, MD, PhD, Division of Gerontology, Geriatrics, and Palliative Care, Department of Medicine, University of Alabama at Birmingham, 933 19th Street South, CH19–218R, Birmingham, AL 35294. rekenned@mail.ad.uab.edu.

Author Contributions: Kennedy: concept, design, data interpretation, drafting, revision and final approval of manuscript. Williams: data analysis, revision and final approval of manuscript. Sawyer: design, data analysis and interpretation, revision and final approval of manuscript. Allman: design, data interpretation, revision and final approval of manuscript. Crowe: concept, design, data interpretation, revision and final approval of manuscript.

Financial Disclosure: Richard E. Kennedy and Courtney P. Williams were supported in part by Awards R01 AG16062 (REK, CPW, PS, MC, RMA), P30AG031054 (PS, RMA), P30AG022838 (MC), and 5UL1 RR025777 (RMA) from the National Institutes of Health.

Sponsor's Role: All sponsors funded the underlying research scope of all respective authors where related aims are applicable to this study, but no specific sponsor was explicitly involved in the design, methods, subject recruitment, data collection, analysis, or preparation of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Veterans Health Administration.

and MMSET; linear and logistic regressions were used to test associations with clinical outcomes of IADL difficulty and verified diagnoses of dementia.

Results—The MMSET showed good internal consistency (Cronbach $\alpha=0.845$), similar to the full MMSE, and strong correlation with the latter (Spearman $\rho=0.694$, $p<.001$). The MMSET explained a similar proportion of IADL difficulty as the full MMSE (coefficient of variation=0.201 and 0.189, respectively). The MMSET was also associated with verified dementia diagnoses (area under the receiver operating characteristic curve=0.73), which was similar to the full MMSE.

Conclusion—The MMSET is a brief, valid measure of cognition in older adults with psychometric properties similar to that of the full MMSE. Because it can be administered over the telephone, further use in epidemiological studies is promising.

Keywords

Mini-Mental State Examination; telephone administration; epidemiological studies; cognitive assessment

The Mini-Mental State Examination (MMSE)¹, a 30-point scale assessing a variety of domains in approximately 5 minutes, is one of the most widely used brief cognitive examinations in the United States². Availability of age- and education-adjusted summary data from the Framingham Study has greatly contributed to acceptance of the MMSE³, although despite widespread use, it requires in-person administration for several items, limiting its use in epidemiological studies. A telephone version of the MMSE would have clear advantages over other instruments because results could be easily compared with those of the full MMSE. Furthermore, research studies that administer both instruments could combine such data longitudinally.

The original MMSE has been modified to remove dependence on visual tasks and make it more amenable to telephone administration^{4,5}. A 22-point version of the MMSE was administered as part of the Adult Lifestyles and Function Interview (ALFI), the ALFI-MMSE, that omitted eight items of the original MMSE that could not be administered without visual cues or assessment.⁴ Validation was performed by telephone administration of the ALFI-MMSE, followed by in-person administration of the MMSE, in 100 consecutive evaluations in a geriatric outpatient program. The two versions showed a high level of overall agreement (Pearson $\rho=0.85$), which was seen regardless of presence or absence of cognitive impairment. The ALFI-MMSE had predictive ability similar to that of the MMSE when using cognitive impairment on a brief neuropsychological screening test as the outcome. A 26-point version, the Telephone MMSE (TMMSE), added a modified three-step command and recall of the individual's telephone number to the ALFI-MMSE.⁵ Validation was performed in 53 consecutive evaluations in individuals followed as part of a longitudinal study of Alzheimer's disease who were administered the MMSE in person followed by telephone administration of the TMMSE. The two versions were highly correlated, with Pearson $\rho=0.88$ ($p<.001$) comparing the TMMSE with the full MMSE and Pearson $\rho=0.88$ ($p<.001$) when comparing the TMMSE with the 22 items in common with the full MMSE. Scores on each of the subscales (orientation to time, orientation to place,

registration, attention, recall) ranged from 0.36 to 0.75 in the two versions, and agreement on individual items (Kappa coefficient) ranged from less than 0.40 to 0.80.

A systematic review of telephone screening instruments for cognitive impairment in older adults noted that the paucity of high-quality studies with a suitable reference standard made recommendations difficult, although based on current evidence, they noted that telephone versions of the MMSE had the advantage of simplicity and good correlation with in-person administration.⁶ Although initial data on the TMMSE have been promising, sample sizes in available validation studies remain small, with few minority participants. Furthermore, validation has been conducted in referrals for clinical evaluation or from other research studies, which may preferentially select individuals who are not representative of the general population. The goal of this study was to describe implementation of the MMSE-Telephone (MMSET) and to validate it against cognitive and clinical measures in a large epidemiological study of African-American and white community-dwelling older adults. It was hypothesized that the telephone MMSE would show good correlation with in-person administration and similar performance in classifying clinical outcomes of dementia and impairment on instrumental activities of daily living (IADLs). By addressing shortcomings of previous studies, it is expected that this validation study will contribute significantly to acceptance of telephone versions of the MMSE in population-based research and community-based programs for older adults that include cognitive assessment.

METHODS

Subjects

All subjects were drawn from the University of Alabama at Birmingham (UAB) Study of Aging II, a longitudinal study of community-dwelling older adults from 17 counties across Alabama designed to examine specific factors predicting mobility decline. Participants in the UAB Study of Aging II were recruited from two prior studies: the UAB Study of Aging, a stratified random sample of community-dwelling Medicare beneficiaries residing in central Alabama designed to be 50% African-American, 50% male, and 50% rural, and the State of Alabama Long Term Needs Assessment Survey (Charting the Course), a sample of community-dwelling adults aged 55 and older recruited from publicly available data banks and screened to represent the state population according to age, race, sex, and geographic region. After obtaining informed consent, trained interviewers conducted in-home baseline interviews between June 2010 and August 2011. Monthly follow-up telephone interviews are on going. Inclusion criteria for the Study of Aging II were aged 75 and older at baseline interview, living independently in the community at baseline, and able to schedule study appointments and answer questions by themselves. Individuals with mild cognitive impairment or mild dementia were not specifically excluded. Additional inclusion criteria for the present analysis were completion of the MMSE and MMSET. The UAB institutional review board approved the study protocol.

Measures

Sociodemographic information and measures of health and cognition were obtained at baseline in-home interviews. A diagnosis of dementia was considered verified if dementia

was reported on a questionnaire that the participant's physician returned or was noted on a hospital discharge summary dated within a year of baseline interview or if the participant reported that a physician had told him or her that he or she had dementia and he or she was taking medication for dementia.

Mini-Mental State Examination

The MMSE is a brief measure of cognition assessing domains of orientation, attention, concentration, memory, visual constructional, and language. Scores range from 0 to 30 points, with lower scores indicating worse performance. The MMSE was administered as part of the baseline in-home interview. For consistency in scoring, the concentration task required participants to spell the word "WORLD" backward rather than subtracting serial sevens, based on pilot data indicating difficulty with numeracy in the study sample.

Mini-Mental State Examination Telephone

The Study of Aging research team created the MMSET to be similar to prior instruments^{4,5} by omitting questions of the MMSE that can only be answered in person: following a 3-step command, reading and repeating a sentence, reading and obeying a command, writing a sentence, and copying intersecting pentagons. The naming task was shortened to one item instead of two, asking the participant to "name the thing you are speaking into as you talk to me (the telephone). As with the MMSE, the concentration task required participants to spell the word "WORLD" backwards. Scores range from 0 to 22 points, with lower scores indicating worse performance. The MMSET was administered during the initial telephone follow-up interview.

Six-Item Screener

Derived from the MMSE, the Six-Item Screener (SIS) is a global measure of cognitive status assessing three-item recall and orientation to year, month, and day of the week. Scores range from 0 to 6, with scores of 4 or fewer indicative of cognitive impairment. The SIS has been validated against clinical diagnoses of dementia and mild cognitive impairment⁷. The SIS was not administered separately in the Study of Aging; scores were obtained from the subset of relevant items on the MMSE or MMSET.

Instrumental Activities of Daily Living

Information on degree of self-reported difficulty in performing IADLs (using the telephone, managing money, preparing meals, doing light housework, shopping, doing heavy housework) was collected in the telephone interview. For each activity, participants were asked whether they had any difficulty performing the task. If they indicated any difficulty, they were asked to rate the level of difficulty as some, a lot, or unable to perform the task. A score of 0 was assigned for no difficulty, 1 for some difficulty, 2 for a lot of difficulty, and 3 for being unable to perform the activity. Difficulty scores were summed for individual tasks to develop IADL composite function scores ranging from 0 to 18, with higher scores representing worse function.

Statistical Analysis

Summary data for the sample consisted of frequencies (percentages) for categorical variables and means and standard deviations for continuous variables. Cronbach alpha was used to determine internal consistency of items for the MMSE and MMSET. Agreement between the MMSE and MMSET was based on total scores and examined using Spearman correlation coefficients because of the skewed distribution of scores. Agreement on individual items common to the MMSE and MMSET was examined using Gwet's first-order agreement coefficient (AC1)⁹ because of high levels of concordance on many items.

To evaluate the clinical utility of each measure, linear regression was used to examine the relationship between degree of IADL difficulty and MMSE and MMSET scores. The coefficient of determination (R^2) was used to measure the proportion of variability in IADL difficulty scores due to variability in MMSE or MMSET scores. Receiver operating characteristic (ROC) curves were also constructed for the MMSE and MMSET using verified dementia diagnosis as the outcome. The ROC curve evaluates sensitivity and specificity of an outcome over a range of cutoff values. The area under the curve (AUC) reflects overall predictive ability, with a value of 0.5 reflecting prediction no better than chance and a value of 1.0 reflecting perfect prediction. All analyses were conducted using JMP version 10.0.2 and SAS version 9.3 (SAS Institute, Inc., Cary, NC).

RESULTS

In-home baseline assessments were completed on 419 eligible older adults. Seventeen participants did not provide follow-up telephone assessments. (Seven withdrew from the study before telephone follow-up could be completed, one had been placed in a nursing home, and nine could not be reached.) The remaining 402 participants are the focus of this article. The 17 excluded participants differed from those who remained in the study in age (mean 84.4 and 81.7, respectively, $p=.02$) and baseline MMSE (mean $\pm$ standard deviation 19.3 $\pm$ 6.2 and 25.8 $\pm$ 4.2, respectively, $p < .001$) but not race, sex, marital status, education, or dementia diagnosis. Mean SIS scores were 4.8 $\pm$ 1.3 for the in-home interview ($n=419$) and 4.8 $\pm$ 1.3 for the telephone interview ($n=402$). Sample characteristics are shown in Table 1. The sample was 58% female and 35% African-American, and 42% had education beyond high school. A dementia diagnosis was verified for 14 participants (11 by physician report and 3 by participant report plus taking medication).

The MMSET showed good internal consistency with the MMSE (Cronbach $\alpha=0.845$ and 0.763, respectively). The MMSET also showed strong correlation with the full MMSE in comparison with the full 30-point MMSE (Spearman $\rho=0.694$, $p<.001$) and with the sum of the 22 items common to both measures (Spearman $\rho=0.688$, $p<.001$). Kappa coefficients of agreement for the 22 individual items ranged from 0.241 (poor agreement) for delayed recall of the word "dollar" to 0.990 (excellent agreement) for orientation to county (Table 2).

As has been reported for the MMSE³, mean MMSET scores were higher with higher education ($p<.001$) (Table 3). The MMSET showed similar performance in percentage of IADL difficulty explained ($R^2=0.201$, $p<.001$) as the MMSE ($R^2=0.189$, $p<.001$). The performance of the MMSET in classifying dementia (AUC=0.73) was similar to that of the

MMSE (AUC=0.70). The AUCs placed both within the “good” range of classification ability.

Restricting analysis to the six items on the SIS, there was good correlation between the SIS in in-home and telephone interviews (Spearman $\rho=0.444$, $p<.001$). Education also highly influenced SIS score ($p<.001$) (Table 3). The two SIS scores explained a significant, albeit smaller, percentage of IADL difficulty ($R^2=0.068$ and 0.123 , respectively, both $p<.001$) than the MMSE and MMSET. The SIS scores had a similar “good” range of classification ability as the MMSE and MMSET (AUC = 0.74 and 0.71, respectively) for verified dementia diagnosis.

DISCUSSION

This study examined a modified version of the MMSE, the MMSET, and demonstrated that the MMSET performed similarly to the MMSE in the UAB Study of Aging II. The MMSET and the MMSE both had good internal consistency, indicating that restricting the MMSET to 22 items that could be administered over the telephone did not affect its performance as a general measure of cognition. There was good to excellent agreement between administrations on most individual items on the MMSET and MMSE, with some concentration and delayed recall items falling into the fair to poor range.

The MMSET was similar to the MMSE in proportion of IADL difficulty explained and in classification of individuals with diagnosed dementia, indicating that restriction of the 30-item MMSE to the 22-item MMSET would have no effect if used in detecting clinically important outcomes. This also indicates that the MMSET is potentially useful as a screening or follow-up measure for epidemiological studies of cognition. In contrast, the SIS performed similarly to the MMSE in classification of dementia, although a smaller proportion of IADL difficulty was explained. The similarity in dementia classification may be due to the small number of dementia cases in the sample; if more individuals with cognitive impairment had been present, differences between the MMSE and SIS might have been observed, suggesting that restriction to six items instead of 22 results in degradation of performance, although the performance observed for the SIS may be acceptable for many applications.

This investigation extends previous validation studies of telephone versions of the MMSE—the ALFI-MMSE and TMMSE—using a large, community-based sample of older adults with concurrent clinical outcomes. The sample size of 402 is more than four times as large as previously reported samples, with good representation of African Americans as well as Caucasians. Because participants were drawn from the community, rather than clinical referrals, these results are more representative of community-based epidemiological studies. Finally, examining the association between the MMSET and IADL difficulty and verified dementia diagnoses establishes the relevance of the MMSET to clinically important outcomes that have not been previously examined.

Despite these considerable strengths, there are weaknesses that must be acknowledged. First, the process of verifying dementia diagnoses may lead to underreporting, particularly with

mild forms of disease, although this would tend to decrease, rather than increase, classification accuracy of the MMSE and MMSET. The use of self-reported IADLs may not be as accurate as clinician ratings. Also, the small number of participants with dementia may limit generalizability to impaired individuals, although prevalence of dementia in this study is consistent with community samples¹⁰. Second, participants were only from Alabama, although this sample is believed to be representative of the population in much of the southeastern United States, where 50% of older African Americans live. Third, participants were limited to Caucasians and African Americans, so results may not be applicable to other racial and ethnic groups, particularly those with limited English language proficiency. Finally, this analysis did not examine the sensitivity of the MMSET to change over time, which is important in choosing an assessment instrument.

In conclusion, the MMSET is a brief, valid measure of cognition in older adults with psychometric properties similar to that of the full MMSE. These results specifically show the feasibility of longitudinal analyses in the UAB Study of Aging using the MMSE and MMSET administered at different time points. Further use of the MMSET in other epidemiological studies requiring cognitive assessment similar to the MMSE is encouraged.

References

1. Folstein MF, Folstein S, McHugh P. 'Mini-mental state' A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975; 12:189–198. [PubMed: 1202204]
2. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: A comprehensive review. *J Am Geriatr Soc.* 1992; 40:922–935. [PubMed: 1512391]
3. Crum R, Anthony J, Bassett S, et al. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA.* 1993; 269:2386–2391. [PubMed: 8479064]
4. Roccaforte WH, Burke WJ, Bayer BL, et al. Validation of a telephone version of the Mini-Mental State Examination. *J Am Geriatr Soc.* 1992; 40:697–702. [PubMed: 1607586]
5. Newkirk LA, Kim JM, Thompson JM, et al. Validation of a 26-point telephone version of the Mini-Mental State Examination. *J Geriatr Psychiatry Neurol.* 2004; 17:81–87. [PubMed: 15157348]
6. Martin-Khan M, Wootton R, Gray L. A systematic review of the reliability of screening for cognitive impairment in older adults by use of standardised assessment tools administered via the telephone. *J Telemed Telecare.* 2010; 16:422–428. [PubMed: 21030488]
7. Callahan CM, Unverzagt FW, Hui SL, et al. Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Med Care.* 2002; 40:771–781. [PubMed: 12218768]
8. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969; 9:179–186. [PubMed: 5349366]
9. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol.* 2008; 61:29–48. [PubMed: 18482474]
10. Jalbert JJ, Daiello LA, Lapane KL. Dementia of the Alzheimer type. *Epidemiol Rev.* 2008; 30:15–34. [PubMed: 18635578]

Table 1

Demographic and Other Characteristics of the Study Sample (N=402)

Characteristic	Value
Age, mean $\pm$ SD	81.6 $\pm$ 4.8
Caucasian, n (%)	263 (65)
Female, n (%)	233 (58)
Education, n (%)	
Elementary school	32 (8)
High school	202 (50)
Some college	126 (3)
College graduate	42 (10%)
Instrumental activity of daily living difficulty score, mean $\pm$ SD (range 0–6)	2.6 $\pm$ 3.9
Dementia diagnosis	14 (3)
Assessment interval, days, mean $\pm$ SD (range 8–56)	21.6 $\pm$ 9.0

SD=standard deviation.

Table 2

Gwet First-Order Agreement Coefficient (AC1) for the 22 Common Items of the Mini-Mental State Examination (MMSE) and MMSE Telephone

Item Description	AC1
Orientation to time	
Day of the week	0.935
Season	0.924
Month	0.942
Day of the month	0.676
Year	0.931
Orientation to place	
County	0.990
Street	0.961
City	0.977
State	0.966
Telephone number	0.963
Immediate recall	
Airplane	0.953
Dollar	0.910
Apple	0.930
Concentration	
D	0.808
L	0.679
R	0.498
O	0.534
W	0.679
Delayed recall	
Airplane	0.579
Dollar	0.241
Apple	0.480
Naming	0.977

Table 3

Mean Scores and Percentage Correct Responses on the Mini-Mental State Examination (MMSE), MMSE Telephone (MMSET), and Six-Item Screener (SIS) According to Education

Education	MMSE	MMSET	In-home SIS	Telephone SIS
	Mean Score (Percentage Correct)			
Elementary school	20.0 (66.6)	14.8 (67.2)	4.1 (68.2)	3.7 (61.5)
High school	25.5 (85.0)	18.1 (82.3)	4.8 (80.0)	4.7 (78.0)
Some college	27.3 (91.1)	19.7 (89.4)	5.0 (84.0)	5.1 (85.4)
College graduate	27.1 (90.4)	19.7 (89.7)	5.0 (82.5)	4.9 (81.3)

For each of the three measures, analysis of variance for educational level is significant at $p < .001$.