UCLA
UCLA Previously Published Works

Title

Validation of a Brief Multi-Dimensional Assessment of Dementia Severity

Permalink

https://escholarship.org/uc/item/6979h7rz

Journal

Journal of the American Geriatrics Society, 69(2)

ISSN

0002-8614

Authors

Peipert, John D
Jennings, Lee A
Romero, Tahmineh
et al.

Publication Date

2021-02-01

DOI

10.1111/jgs.16930

Peer reviewed

BRIEF REPORT

Validation of a Brief Multi-Dimensional Assessment of Dementia Severity

John D. Peipert, PhD, * Lee A. Jennings, MD, MSHS,† Tahmineh Romero, MS,‡*
Ron D. Hays, PhD,‡ Neil S. Wenger, MD,‡ Emmett Keeler, PhD,§ and
David B. Reuben, MD¶

BACKGROUND/OBJECTIVES: Briefer measures of symptoms and functional limitations may reduce assessment burden and facilitate monitoring populations of persons with dementia (PWD).

DESIGN: Prospective follow-up study.

SETTING: University-based dementia care management program.

PARTICIPANTS: 1,091 PWD.

MEASUREMENTS: We assessed cognition (Mini Mental State Examination (MMSE)—11 tasks), neuropsychiatric symptom severity (Neuropsychiatric Inventory Questionnaire Severity Scale (NPIQ-S)—12 items), and functional ability (Activities of Daily Living (ADL)—6 items; Functional Activities Questionnaire (FAQ)—10 items). Item response theory was used to select subsets of items by identifying low item discrimination (<1.50), poor item fit (χ^2), local dependence (LD), and with difficulty similar to other items. We estimated correlations between original and shorter scales and compared their associations with mortality. We added two symptoms (trouble swallowing, coughing when eating) reflecting late-stage dementia complications, created a multi-dimensional dementia assessment composite, and examined its association with mortality.

RESULTS: Five MMSE tasks were eliminated: two with low discrimination, two with difficulty similar to other items, and one with poor fit. The remaining tasks were correlated with the full MMSE at $r = 0.82$. We retained three ADLs that were correlated with the total ADL set at $r = 0.95$ and kept five FAQ items that were not LD (correlation with full FAQ, $r = 0.97$). Associations with mortality were similar between the longer and shorter scales. A higher score on the composite (range 0–100) indicates worse dementia impact and was associated with mortality (hazard ratio (HR) per scale point: 1.03 (1.02–1.04)).

CONCLUSION: These brief assessments and dementia composite may reduce administration time while preserving validity. **J Am Geriatr Soc 00:1-5, 2020.**

Keywords: dementia; Alzheimer's; behavioral symptoms; cognitive assessment

Abbreviations

ADC	Alzheimer's and dementia care
ADL	Activities of daily living
ADRC	Alzheimer's Disease Research Center
FAQ	Functional activities questionnaire
GRM	Graded response model
IQR	Interquartile range
IRT	Item response theory
LD	Local dependence
MMSE	Mini Mental State Examination
MOCA	Montreal cognitive assessment
NACC	National Alzheimer's coordinating center
NIA	National Institute on Aging
NPIQ-S	Neuropsychiatric Inventory Questionnaire Severity Scale
PWD	Persons with dementia
UCLA	University of California, Los Angeles

INTRODUCTION

Persons with dementia (PWD) experience cognitive, functional, and neuropsychiatric impairments,

including depression,[1] apathy,[2] hallucinations,[1] and language impairment.[3] Most PWD exhibit more than one symptom simultaneously.[2,4] Instruments used to measure dementia symptoms and complications often exceed 10 questions each and usually measure only one dimension (e.g., cognition). Briefer versions of these instruments may reduce time spent conducting assessments.

We conducted psychometric analyses of the Mini Mental State Examination (MMSE; 30 items across 11 tasks), Neuropsychiatric Inventory Questionnaire Severity Scale (NPIQ-S; 12 items), and functional ability, indicated by Activities of Daily Living (ADL; 6 items) and the Functional Activities Questionnaire (FAQ; 10 items; represents instrumental ADLs (IADLs)), aiming to generate shorter versions of each. In addition, we assessed the impact of two symptoms of late stage dementia (trouble swallowing and coughing when eating). Finally, we created a short, multidimensional composite scale to assess dementia. As tests of validity, we hypothesized that PWD with more severe ratings on the shorter scales and composite would have higher likelihood of mortality.

METHODS

Study Sample

Participants were 1,091 dyads of PWD and their primary caregivers enrolled in the UCLA ADC program between July 2012 and December 2014.[6] The ADC serves community-dwelling patients diagnosed with any type of dementia who are referred by a UCLA provider. Caregivers complete standardized clinical instruments during routine care for the PWD in the program. The data for this study were taken from surveys at time of program entry. PWD were followed for a median of 1.9 (IQR: 1.1–2.5) years to examine mortality. The UCLA IRB approved the study protocol (#13-001480-AM-00014).

Survey and Measures

The study measured five areas that represent manifestations and complications of dementia: cognition, behavioral symptoms, functional status (ADLs and IADLs), and swallowing difficulties. The MMSE was used to measure cognitive functioning[5] and consists of 11 cognitive tasks (e.g., delayed recall of three words), the correct performance of which is scored with one or more points. The total MMSE score ranges from 0 to 30, with higher scores indicating better cognitive functioning. Behavioral symptoms were assessed using the NPIQ-S.[6] It was completed by caregivers and assesses experience of 12 symptoms, rated as "not present," "mild," "moderate," or "severe." We captured two aspects of functional ability. First, ADLs were assessed using a 6-item (higher is better) basic activities of daily living scale (can perform/cannot perform).[7] Second, the FAQ,[8] a measure of IADLs, was also completed by caregivers. The FAQ asks how often the PWD needed help over the past 4 weeks with 10 activities rated as "normal," "has difficulty, but does by self," "requires assistance," or "dependent." Finally, two questions were asked to assess medical complications that have demonstrated a strong association with late, terminal stage dementia.[9] These included whether the

PWD had problems swallowing (yes/no) or coughed while eating (yes/no). Additional details of each measure are given in the Supplementary Materials.

Statistical Analyses

Analyses aimed to: (1) identify subsets of items from the MMSE, FAQ, ADL, and NPIQ-S scales (item reduction); (2) examine reliability and validity of the identified item subsets compared to original versions; and (3) create a multidimensional composite. For item reduction analyses, we used item response theory (IRT) graded response model parameters. In these analyses, every item in the NPIQ-S, ADL and the FAQ was entered, and for MMSE, the score for each cognitive task was entered. We used several statistics from the IRT models to select subsets of items from each original scale, including evaluations of each item's ability to discriminate between patients with higher and lower symptom impact or functional impairment and whether each item provided unique information in comparison to other items. The details of these analyses are given in the Supplementary Materials.

We compared the internal consistency reliability of the item subsets with the original scale versions using coefficient alpha interpretations: 0.70 and higher = acceptable; 0.80 and higher = good; 0.90 and higher = excellent. We assessed the validity of the shorter versions of each scale using several analyses. First, we calculated Pearson's correlation coefficients correspondence between the shortened and original scale versions. Next, we tested the shorter scales' predictive validity by comparing their associations with mortality with the longer versions. Validity tests for the shorter scales and composite using mortality were based on the hypothesis that PWD with more severe cognitive impairment, neurocognitive symptoms, and functional impairment would have a higher likelihood of death. If the shorter scales' associations and predictive ability were similar to the longer versions, we took this as evidence that they were as valid as the longer versions. To conduct these analyses, we fit a series of Cox proportional hazards models using the proc phreg procedure in SAS. First, for each short and long version of the scales, we fit separate Cox models and compared hazard ratios (HRs), regression betas, and Harrell's c-statistic[10] for the short and full scales.

We created a composite representing a brief multidimensional dementia assessment. This composite was calculated by transforming each item from the shortened scales and two clinical symptoms (problems swallowing and coughing while eating) linearly to a 0–100 possible range scale, with higher scores indicating more severe impairment. Then, the mean of retained items was taken to generate a final score. More details about this scoring algorithm are given in the Supplementary Materials. We compared the mean composite scores across key demographic groups, including sex, age, education, and language spoken.

With similar rationale as the validity analyses using mortality described above, we examined the composite's association with mortality after adjusting for other clinically relevant covariates, including age, sex, history of heart attack, history of heart failure, history of atrial fibrillation, history of diabetes mellitus, high blood pressure, high cholesterol, Parkinson's disease, history of stroke, and

depression. If the composite is a valid dementia assessment, PWD with higher scores should have higher probability of mortality, even when controlling for relevant risk factors. Each covariate was simultaneously entered into a multivariable Cox regression model without the composite. Then, those with effects significant at $P < .05$ and well-established risk factors for mortality (i.e., prior heart attack and heart failure) were retained for a second model in which they were entered simultaneously with the composite.

RESULTS

Participant Characteristics

Characteristics of PWD are provided in Table 1.

Scale Reduction

A general overview of IRT analyses to select a smaller set of items from each scale is given here, and additional detail is provided in the Supplementary Materials (Supplementary Table S1). Six tasks from the MMSE were retained after

Table 1. Persons with Dementia Characteristics

Age, mean (SD, range)	82 (9, 40–101)
Sex, % (n)	
Female	66% (720)
Male	34% (371)
Ethnicity, % (n)	
Hispanic/Latino	12% (124)
Not Hispanic/Latino	88% (879)
Race, % (n)	
White	75% (808)
African American	9% (102)
Asian	6% (70)
Other	<1% (16)
Level of education, % (n)	
High school or less	34% (355)
Some college	22% (226)
College graduate	20% (215)
Graduate school	24% (255)
Language, % (n)	
English	83% (868)
Spanish	8% (82)
Other	9% (102)
Caregiver lives with person with dementia, % (n)	
Yes	55% (595)
No	34% (369)
Missing	12% (127)
Diagnosis, % (n)	
Dementia not specified	43% (465)
Alzheimer's dementia	38% (413)
Mixed dementia	8% (88)
Vascular dementia	4% (44)
Lewy body dementia	3% (38)
Frontotemporal dementia	1% (17)
Parkinson's disease	1% (16)
Missing	<1% (10)
MMSE Score, mean (SD, range)	17 (7, 0–30)
NPIQ-S Score, mean (SD, range)	10 (7, 0–35)
FAQ Score, mean (SD, range)	21 (8, 0–30)

eliminating two for low discrimination, one for poor item fit, and two with difficulties similar to other items. The retained tasks include: naming a pencil and watch (language); reading and obeying the command (language); copying a geometric design (visuospatial); naming the year, season, date, day, and month (orientation); repeating three objects named by the test administrator (registration); naming the three objects from the registration task (recall). Nine of the 12 NPIQ-S items were omitted for low discrimination values, and we retained the agitation/aggression, disinhibition, and irritability items. We removed two FAQ items for poor item fit and three for local dependence, leaving the following five items: assembling tax records, business affairs or papers; shopping alone for clothes, household necessities or groceries; preparing a balanced meal; paying attention/understanding/discussing TV program, book or magazine; and remembering appointments, family occasions, holidays, or medications. Finally, we omitted two ADL items for poor item fit and one for similar difficulty to other items, leaving: getting from bed to chair, getting dressed, and bathing or showering items.

Reliability of Items from the Full Versions of the Scales

For each scale, the selected item set exhibited at least acceptable reliability (>0.70; MMSE, NPIQ-S) or good reliability (>0.80; FAQ, ADL). The reliability estimates are provided in Supplementary Table S2.

Correlations Between Items from the Scales and Full Versions of the Scales

Each of the reduced items of the scales was correlated with its corresponding full version. The highest correlations were between the shortened and full FAQ ($r = 0.97$, $R^2 = 0.94$, reduced item set explained 94% of the variance of the full version) and the ADLs ($r = 0.95$, $R^2 = 0.90$). The items drawn from the MMSE and full MMSE were correlated at $r = 0.82$ ($R^2 = 0.67$). Finally, items from the NPIQ-S and full NPIQ-S was correlated at $r = 0.80$ ($R^2 = 0.64$).

Associations with Mortality

The maximum follow-up time for study participants was approximately 4 years, and approximately 59% were still alive at this time. The Supplementary Materials contains additional information about the survival models used in this analysis. The HRs, regression betas, and Harrell's c-statistics for associations between both the full and shortened scales with time to mortality are shown in Table 2. Comparing across the full versions and items selected from each scale, the HRs were largely similar. For each scale, the Harrell's c-statistic was very similar across the short and long versions, indicating similar associations and predictive ability between each version and mortality, suggesting that the shorter versions of the scales are as valid as the longer versions.
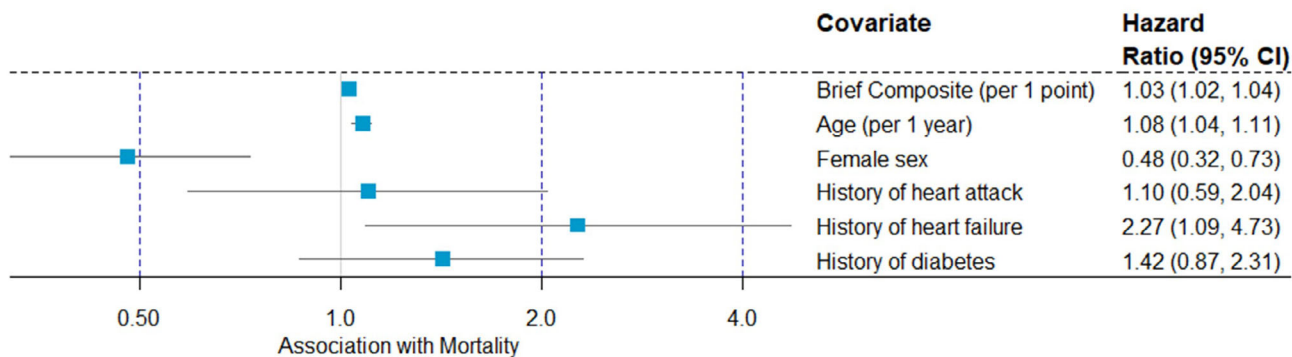
Brief Multidimensional Dementia Assessment

We created a composite from all the items or tasks selected from the MMSE, NPIQ-S, FAQ, ADLs, and the two medical symptoms. The composite had a mean of 46 (SD = 13), an observed minimum of 0, a 25th percentile score of 31,

Table 2. Shortened (Selected Items) and Full Scales' Associations with Time to Mortality

	Selected Items				Full Scale Score			
	Hazard Ratio (95% CI)	Beta (SE)	*P*-value	Harrell's C-Statistic	Hazard Ratio (95% CI)	Beta (SE)	*P*-value	Harrell's C-Statistic
Mini Mental State Examination	1.08 (1.01–1.14)	0.07 (0.03)	0.016	0.60	1.04 (1.01–1.07)	0.04 (0.01)	0.007	0.61
Functional Assessment Questionnaire	1.09 (1.03–1.15)	0.08 (0.02)	0.003	0.59	1.05 (1.02–1.08)	0.05 (0.01)	<0.001	0.61
Neuropsychiatric Inventory Questionnaire-Severity	1.12 (1.05–1.20)	0.12 (0.04)	<0.001	0.58	1.05 (1.02–1.07)	0.04 (0.01)	<0.001	0.60
Activities of daily living	1.54 (1.33–1.79)	0.43 (0.08)	<0.001	0.68	1.28 (1.18–1.38)	0.25 (0.04)	<0.001	0.68

Covariate	Hazard Ratio (95% CI)
Brief Composite (per 1 point)	1.03 (1.02, 1.04)
Age (per 1 year)	1.08 (1.04, 1.11)
Female sex	0.48 (0.32, 0.73)
History of heart attack	1.10 (0.59, 2.04)
History of heart failure	2.27 (1.09, 4.73)
History of diabetes	1.42 (0.87, 2.31)

Harrell's c-statistic = 0.75

Figure 1. Multivariable cox model of the dementia severity composite's association with mortality adjusted for clinically relevant covariates.

50th percentile score (median) of 46, 75th percentile score of 61, and a maximum observed score of 100. Composite scores tended to vary by demographic groups, especially by age and education level (Supplementary Table S3). Adjusting for clinical covariates associated with time to mortality (age, female sex, history of heart attack, history of heart failure, and history of diabetes mellitus), the composite (per 1 point change) was still independently associated with mortality (HR = 1.03, 95% CI = 1.02–1.04; Figure 1). The Harrell's c-statistic for this model was 0.75. These results provide evidence that the composite is a valid dementia assessment.

DISCUSSION

This study identified subsets items from assessments commonly used in dementia research and clinical care, including the MMSE (cognition), NPIQ-S (neuropsychiatric symptoms), ADLs (functional ability), FAQ (functional ability), as well as two additional indicators of late stage disease related to eating problems. In doing so, we reduced the assessment for these common measures significantly (from 41 to 19 tasks or items). Taken as a whole, these 19 tasks or items capture both a diversity of constructs (cognition,

function, and complications) and a range from unimpaired to severely impaired.

Results of our study are consistent with some prior work but differed in some ways from others. For example, in a review of IRT applied to the MMSE in PWD, date orientation, serial 7s, and recall were the most difficult tasks.[11] In our analysis, each of these tasks were also among the most difficult. However, results of another IRT analysis of the ADLs, IADLs similar to FAQ, and NPIQ were less consistent with our own analyses.[12-14] In addition, previous research has found that eating problems similar to the two symptoms we included (trouble swallowing, coughing while eating) were very common in the last months of life and had a very strong association with mortality.[9]

While the reduced sets of selected items and the composite are not measures of medical comorbidity or complications of dementia, they may have value for outcomes tracking and related research, similar to other measures currently implemented by the NIA Alzheimer's Disease Research Centers (ADRCs).[15] In addition, though none of the tools should be used for prognosis, they may be useful as brief, gross screeners of decline leading toward death. However, we do note that the longer versions of the scales are more appropriate for individual patient management due to the inclusion of more content (e.g., the selected

NPIQ-S items only cover a narrow range of symptoms associated with severe disease), are more reliable and would be more appropriate as a multi-dimensional guide to care planning. In terms of research, these shorter item sets and the dementia composite will offer advantages in their brevity and may be particularly appropriate for characterizing patient populations and for use in registries. The advantage of the composite over existing instruments is that it brings in more dimensions (cognition, behavioral symptoms, functional ability, and clinical complications) and includes an objective measure of cognition.

This study's limitations must be considered. First, this research is preliminary and additional validation is needed. Specifically, we recommend comparison of scores across patients who have been staged on the CDR scale or other staging scales,[16] examining associations and ability to predict non-mortality outcomes, and comparison of performance across demographic and clinical groups that vary in terms of performance. Second, these data were collected from a single, urban, academic health system, and the results may not generalize to the national dementia patient population. Finally, for both the short and long scales, Harrell's c-statistics were relatively low. These results are intended to reflect the scales' expected relationships with mortality to evidence their validity. However, we would caution against using these measures for prognostic purposes.

In conclusion, using a reduced number of items selected from the MMSE, NPIQ-S, FAQ, and ADLs exhibited reliability and validity, and they may be suitable for clinical outcomes tracking and research. Moreover, combining these with two questions about swallowing complications provides a brief, multi-dimensional assessment of dementia.

ACKNOWLEDGMENTS

REFERENCES

1. Auning E, Rongve A, Fladby T, et al. Early and presenting symptoms of dementia with Lewy bodies. Dement Geriatr Cogn Disord. 2011;32(3):202-208.
2. O'Brien JT, Thomas A. Vascular dementia. Lancet. 2015;386(10004):1698-1706.
3. Ferris SH, Farlow M. Language impairment in Alzheimer's disease and benefits of acetylcholinesterase inhibitors. Clin Interv Aging. 2013;8:1007-1014.
4. Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S. Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. JAMA. 2002;288(12):1475-1483.
5. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-198.
6. Kaufer DI, Cummings JL, Ketchel P, et al. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. J Neuropsychiatry Clin Neurosci. 2000;12(2):233-239.
7. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist. 1970;10(1):20-30.
8. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol. 1982;37(3):323-329.
9. Mitchell SL, Teno JM, Kiely DK, et al. The clinical course of advanced dementia. N Engl J Med. 2009;361(16):1529-1538.
10. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247(18):2543-2546.
11. McGrory S, Doherty JM, Austin EJ, Starr JM, Shenkin SD. Item response theory analysis of cognitive tests in people with dementia: a systematic review. BMC Psychiatry. 2014;14(1):47.
12. Fuller JT, Choudhury TK, Lowe DA, Balsis S. Alzheimer's disease neuroimaging Initiative. Hallucinations and delusions signal Alzheimer's associated cognitive dysfunction more strongly compared to other neuropsychiatric symptoms. J Gerontol: B Psychol Sci Soc Sci. 2020;75:1894–1904.
13. McGrory S, Shenkin SD, Austin EJ, Starr JM. Lawton IADL scale in dementia: can item response theory make it more informative? Age Ageing. 2013;43(4):491-495.
14. Fieo R, Zahodne L, Tang MX, Manly JJ, Cohen R, Stern Y. The historical progression from ADL scrutiny to IADL to advanced ADL: assessing functional status in the earliest stages of dementia. J Gerontol: SerA. 2017;73(12):1695-1700.
15. National Institute on Aging. Alzheimer's Disease Research Centers. National Institutes of Health; 2019. https://www.nia.nih.gov/health/alzheimers-disease-research-centers. Accessed January 7, 2020.
16. Burke WJ, Miller JP, Rubin EH, et al. Reliability of the Washington University clinical dementia rating. Arch Neurol. 1988;45(1):31-32.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplementary Appendix S1 Supporting information