ORIGINAL CONTRIBUTION

A Validation Study of the Dementia Questionnaire

Claudia Kawas, MD; James Segal; Walter F. Stewart, PhD, MPH; Maria Corrada, MSc; Leon J. Thal, MD

Objective: To determine the validity of the Dementia Questionnaire (a semistructured informant interview) for the diagnosis of dementia.

Design: Comparison of dementia status determined by a telephone-administered informant questionnaire with the criterion standard of clinical diagnosis established by examination and laboratory studies.

Setting: Gerontology Research Center, the Baltimore Longitudinal Study of Aging.

Subjects: Volunteer cohort of 42 men and 32 women aged 68 to 97 years. Subjects were selected from strata defined by Blessed Information Memory Concentration Test scores, with oversampling of borderline scores (3 to 10).

Main Outcome Measures: Sensitivity and specificity of the Dementia Questionnaire in comparison with the criterion standard of clinical diagnosis.

Secondary Outcome Measure: Interrater reliability (κ coefficient).

Results: Sensitivity and specificity for dementia were 100% and 90%, respectively. Most false-positive findings were from subjects with cognitive impairment that did not meet criteria for dementia (*Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*. Interrater reliability was high (κ=0.83).

Conclusion: The Dementia Questionnaire can be used effectively in research studies to screen for dementia.

(Arch Neurol. 1994;51:901-906)

From the Department of Neurology, The Johns Hopkins University School of Medicine (Dr Kawas and Mr Segal and Ms Corrada), and the Department of Epidemiology, The Johns Hopkins University School of Hygiene and Public Health (Dr Stewart), Baltimore, Md; and the Department of Neurosciences, University of California—San Diego School of Medicine (Dr Thal).

ASCERTAINING cases of dementia in population-based and family history studies presents a number of challenges. Often, the patient either has died or is otherwise unavailable for examination, and medical record documentation is limited. Moreover, diagnosis of dementia is based, in part, on a decline in intellectual abilities rather than simply on cognitive status at the time of an evaluation.[1] As such, mild cases in particular are likely to be underascertained when classification is solely based on clinical examination and psychometric results.

Questions regarding a change or decline in abilities can often be answered by friends or relatives of the subject. For these reasons, informant reports are routinely used in clinical practice in conjunction with cognitive screening instruments, such as the Blessed Information Memory Concentration (IMC) Test[2] and the Mini-Mental State Examination.[3] Informant reports can also be useful in epidemiologic studies to screen for cases of dementia. However, knowledge is limited about the validity of such instruments for screening. For example, the Informant Questionnaire on Cognitive Decline in the Elderly[4] is a validated self-administered informant questionnaire administered by mail to identify subjects who have experienced cognitive loss. Data collected using this ques-

See Subjects and Methods on next page

SUBJECTS AND METHODS

Subjects for the validation study were selected from participants evaluated during 1990 and 1991 as part of the Baltimore Longitudinal Study of Aging (BLSA). The BLSA is a longitudinal study of normal aging in volunteer recruits that began in 1958. While enrollment was initially limited to men, recruitment of women volunteers began in 1978. Most BLSA participants have more than 14 years of education. Participants return to Baltimore every 2 years for a multidisciplinary examination that includes the Blessed IMC Test. Most BLSA subjects are normal, with low Blessed IMC error scores. To ensure an adequate sample of subjects with high Blessed IMC scores, subjects with scores of 3 or more errors (of a possible 34) were oversampled. Subjects were selected without knowledge of their dementia status. Forty-six percent (34 subjects) of the sample had Blessed IMC scores of 0 to 2, 38% (28 subjects) had scores of 3 to 10, and 16% (12 subjects) had scores of more than 10 errors. In contrast, the entire BLSA cohort older than 65 years has the following distribution: 81% (384 subjects) have Blessed IMC scores of 0 to 2 errors, 18% (85 subjects) have scores of 3 to 10, and 1% (seven subjects) have more than 10 errors on their most recent Blessed IMC Test. The sample for this study included 42 men and 32 women between the ages of 68 and 97 years.

Each subject had a standardized neurological examination and neuropsychological testing, in addition to the usual BLSA protocol.[9] Neuropsychological evaluation included the Blessed IMC Test, Mini-Mental State Examination, Immediate and Delayed Cued Recall,[10] Boston Naming Test,[11] Controlled Verbal Fluency (fruits, animals, vegetables, /F/, /A/, /S/),[12] Trail Making Test parts A and B,[13] Clock Drawing and other constructions, Center for Epidemiologic Studies Depression scale,[14] and Pfeffer Functional Activities Questionnaire.[15] The evaluation of all clinical and neuropsychological information was used to classify subjects as (1) normal; (2) normal with mild cognitive loss making early dementia a possibility (suspects); or (3) meeting criteria for dementia (*DSM-III-R*) and/or AD (NINCDS-ADRDA).[16] This decision served as the criterion standard (gold-standard) diagnosis. Six of the subjects (five controls, one with AD) subsequently had verification of diagnosis by autopsy.

After each examination, permission was requested to contact a relative or friend for information about the subject's memory, medical history, and functional abilities. The subject provided the name and phone number of an appropriate informant. Two subjects declined to participate. A trained research assistant, "blinded" to the results of the clinical evaluation, administered the DQ by telephone to 35 spouses, 24 adult children, and 15 other informants. Approximately half of the informants did not live in the same location as the subject. Informants for subjects were contacted within 12 months of examination. Each interview required approximately 20 minutes to complete. The DQ is divided into six parts: (1) memory, (2) language and expression, (3) daily functioning, (4) other medical problems relevant to dementia, (5) medical contacts, and (6) questions relating to patient and family awareness of the problem. It includes questions designed to rule out other possible causes of dementia, such as alcohol abuse, Parkinson's disease, and depressive disorders. It also dates the onset of initial symptoms.

Initial pilot use of the DQ made it apparent that certain probing questions were vital for adequate interpretation. The most important of these involved history of stroke. The interviewer asked for the date of the stroke, whether there were speech or motor problems, and whether there were memory problems before and after the stroke. When asking about the progression of memory loss, the choice of "no change" was added to the choices of "steady downhill progression" and "abrupt declines." If the surrogate reported trouble with money, the interviewer asked if the subject could balance the checkbook, write checks, make change, or figure out how much to tip. Problems with dressing, feeding, or household tasks were explored by asking if the difficulties were due to cognitive or physical problems. This information was recorded extensively on the form.

Using only information obtained from the DQ and applying *DSM-III-R* and/or NINCDS-ADRDA criteria, two of us (C.K. and L.T.) independently diagnosed each subject as normal, suspect, or having dementia and indicated the specific type of dementia. The neurologists were blinded to the subject's identity. In addition to the clinicians' ratings, a computer-applied algorithm compatible with *DSM-III-R*/NINCDS-ADRDA criteria was constructed using specific questions on the DQ (**Figure 1**). For example, if question 1 or 6 was answered "yes," or if question 48 reported forgetfulness of dates or names, the first stage of the algorithm was considered positive. The remaining stages were similarly examined. Question 48 ("What was noticed?") was open ended. Responses were recorded as free field information if the impairment was relevant to the diagnosis by *DSM-III-R* criteria and were coded in reference to one or more stages of the algorithm. For example, if the first change noticed was "forgets names, things told, etc," stage I (memory loss) was considered positive. If the subject was reported to "get lost," stage II of the algorithm (another sphere of cognition) was considered positive for visuospatial problems. Difficulty on the job, driving, managing money, etc, were indicative of functional impairment (stage III).

The κ coefficient[17] was estimated as a measure of interrater reliability between the two neurologists defining subjects as demented or not, and separately as normal, suspect, or demented. Sensitivity and specificity[18] were estimated for dementia vs no dementia (combined normal and suspect). The clinical and DQ diagnoses were also compared using normal, suspect, and dementia categories to better characterize the discordant decisions.

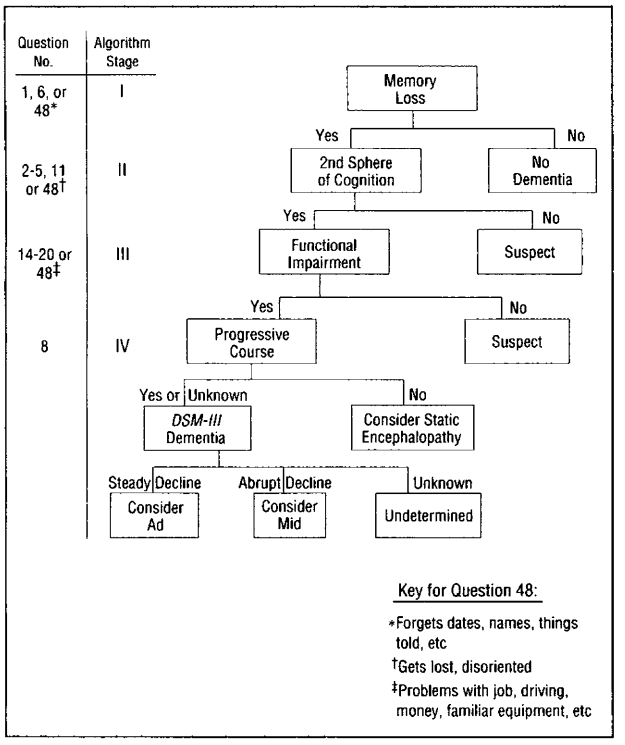

Figure 1. *Diagnostic algorithm and relevant Dementia Questionnaire questions. MID indicates multi-infarct dementia;* DSM-III, Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition; *and AD, Alzheimer's disease.*

tionnaire is insufficient to determine dementia status using the *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*, (*DSM-III-R*), criteria. Other mailed questionnaires designed for specific dementia diagnoses have been reported to be highly sensitive (0.93) but have low specificity (0.43) for Alzheimer's disease (AD).[5] The quality of information elicited by a mailed questionnaire may differ from that obtained by a skilled interviewer who can probe for information and interpret responses.

As an alternative to mailed questionnaires, some investigators have used a 40-minute structured telephone interview administered by physicians with clinical experience in dementia. Sensitivity for probable AD was 88% and specificity was 88% in a series of patients with postmortem verification of the diagnosis.[6] However, most of the AD subjects had severe dementia, making comparison with normal controls easier, and therefore agreement more likely. In addition, this instrument is labor-intensive for the physician and has not been tested in a broader population sample in which the prevalence of dementia is likely to be lower and milder cases are likely to be more common.

As a cost-efficient alternative, we decided to validate a brief structured interview, the Dementia Questionnaire (DQ),[7] administered by a research assistant to a volunteer cohort of normal subjects or those who had mild dementia (**Table 1**). The DQ is a structured informant interview originally developed by Silverman et al[7] to diagnose dementia among relatives of AD probands in family history studies. Interinformant reliability of the DQ is 0.91[7] and interrater reliability is 0.94.[8] To date, however, DQ classification of dementia diagnoses has not been validated against a clinician's assessment. In this article, we describe the results of a validity study comparing DQ diagnosis with a full clinical assessment.

RESULTS

CLINICAL EXAMINATION

The distribution of Blessed IMC scores for all subjects is shown in **Figure 2**. Nineteen subjects were diagnosed by clinical examination, as having AD, one subject had multi-infarct dementia, and 54 subjects were classified as normal controls. For 18 of the 54 controls, cognitive or functional impairment was identified, but the subjects did not meet criteria for dementia (suspects).

DEMENTIA QUESTIONNAIRE

Interrater reliability in judging the presence or absence of dementia from the DQ was relatively high (κ=0.83). All of the discordant decisions were found to occur when one of the two raters (neurologists) classified the subject as suspect (**Table 2**). Six subjects were classified as demented by one rater and suspect by the other; four subjects were classified as normal by one rater and suspect by the other.

To estimate specificity of the DQ for detecting dementia in the BLSA population older than 65 years, we considered the distribution of the BLSA cohort by their most recent Blessed IMC score. During the period this sample was selected, 81% of BLSA subjects in this age group had Blessed IMC scores of 0 to 2 errors, and 19% had scores of 3 or more errors. These proportions were used as weights for calculating specificity. The overall specificity for the BLSA cohort was estimated as 91%. Virtually all errors in DQ diagnoses were for subjects with Blessed IMC scores between 4 and 6. **Table 3** shows that most subjects who were misclassified as having dementia (eight by rater 1 and five by rater 2) when using the DQ were classified as being cognitively suspect using the criterion standard of clinical diagnosis. In part, these discordant decisions may reflect uncertainty in the clinical diagnosis, especially in identifying cases in the early stages of dementia.

In general, the sensitivity (95%) and specificity (92%) of the computer-applied algorithm were similar to the clinicians'. The lower sensitivity of the algorithm compared with that of the clinician was primarily due to ancillary information from the DQ that was not coded.

Table 1. Dementia Questionnaire

	Yes	No	Don't Know	Date
Memory				
Did (does) the subject have any problems with				
1. Memory				
2. Remembering people's names				
3. Recognizing familiar faces				
4. Finding way about indoors				
5. Finding way on familiar streets				
6. Remembering a short list of items				
7. Did trouble with memory begin suddenly _____ or slowly _____				
8. Has the course of the memory problems been a steady downhill progression _____ or have there been abrupt declines _____				
9. Ever see a doctor for memory problems				
10. If yes, what was the cause given _____				
Expression				
11. Ever have trouble finding the right word or expressing self				
12. Talking less over time				
13. Tendency to dwell in the past				
Daily Functioning				
14. Trouble with household tasks				
15. Handling money				
16. Grasping situations or explanations				
17. Difficulty at work (check if N/A _____)				
Age retired _____ Date retired _____				
Date significant change in work status _____				
18. Trouble dressing or caring for self				
19. Trouble feeding self				
20. Controlling bladder and bowels				
21. Agitation and nervousness				
Other problems				
22. High blood pressure				
23. Stroke				
24. More than one (1) stroke				
25. Is one side of the body weaker than the other side				
26. Parkinson's disease (tremors, shuffling gait, rigidity of limbs)				
27. Injury to head resulting in loss of consciousness for more than a second or two				
28. Seizure or fits				
29. Syphilis				
30. Diabetes				
31. Drinking problem (if alcoholism suspected, explore further SADS Sxs)				
32. Did memory problems coincide with drinking				
33. Ever depressed or sad for 2 weeks or more				
34. If yes, ever seek treatment				
35. Ever very high, euphoric, top of the world				
36. If yes, ever seek treatment				
37. Ever seek psychiatric or psychological help for any reason				
38. If yes, ever hospitalized for psychiatric illness				
Where? _____				
39. Down's syndrome				
40. Other medical problems we have not talked about _____				

Table 1. Dementia Questionnaire (cont)

	Yes	No	Don't Know	Date
Medical Contacts				
41. Name and address of first doctor seen for problems:				
42. Ever receive medications				
43. A neurological or psychiatric examination				
44. CAT scan of head				
45. Ever in an institution (Nursing Home)				
Where?				
46. What was diagnosis given for problems				
Recognition of Problem				
47. Who was first person to notice something wrong?				
48. What was noticed?				
49. When was the last time (the subject) seemed to be really well or his/her old self?				

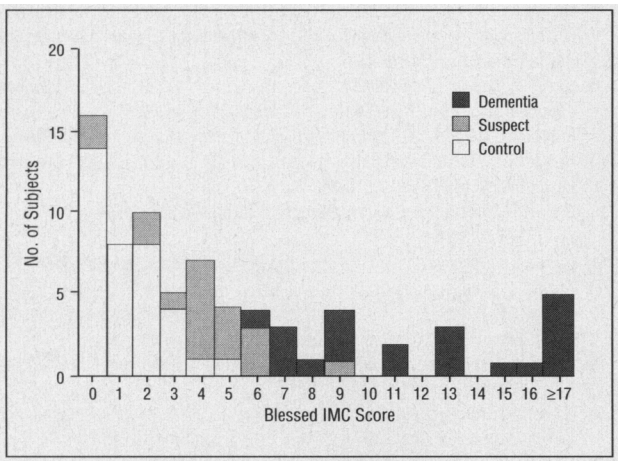

Figure 2. *Distribution of Blessed Information Memory Concentration (IMC) Test scores by clinical diagnosis in Baltimore Longitudinal Study of Aging subjects (N=74).*

Table 2. Interrater Reliability in Assigning a Dementia Questionnaire (DQ) Rating for Dementia Between Raters 1 and 2*

DQ Rating, Rater 2	DQ Rating, Rater 1: Normal	Suspect	Dementia	Total
Normal	34	1	0	35
Suspect	3	6	4	13
Dementia	0	2	24	26
Total	37	9	28	74

*κ *(dementia vs no dementia)=0.83.*

Table 3. Comparison of Clinical Diagnosis to the Dementia Questionnaire (DQ) Rating Using Three Diagnostic Categories: Normal, Suspect, and Dementia

DQ Rating	Clinical Diagnosis, % (No.): Normal	Suspect	Dementia	Total, No.
Normal				
Rater 1	92 (33)	22 (4)	0 (0)	37
Rater 2	89 (32)	17 (3)	0 (0)	35
Suspect				
Rater 1	8 (3)	33 (6)	0 (0)	9
Rater 2	8 (3)	56 (10)	0 (0)	13
Dementia				
Rater 1	0 (0)	44 (8)	100 (20)	28
Rater 2	3 (1)	28 (5)	100 (20)	26
Total	100 (36)	100 (18)	100 (20)	74

COMMENT

This study demonstrated that all clinical cases of dementia were ascertained by the DQ (sensitivity, 100%), but that several subjects were apparently falsely designated as cases of dementia. While specificity of diagnosis was moderate in this study, we believe it is underestimated. Individuals with dementia with higher education may compensate for loss of cognitive abilities through collateral strategies; hence, these individuals may score higher on simple mental status tests in the early stages of the disease, giving rise to misdiagnosis. In contrast, informants are aware of cognitive and functional decline con-

sistent with dementia. Continued follow-up of the discordant cases will be revealing. It is likely that many of these subjects may meet criteria for dementia in future evaluations.

Although in this study the DQ was used to independently determine dementia status, the questionnaire also provides a means to identify other sources of information and hospital-physician records, thereby further improving diagnostic accuracy. In our studies, we generally use the DQ in combination with other measures and information. Nevertheless, these results suggest that the DQ can be used alone to ascertain diagnosis with reasonable accuracy.

Our experience with the DQ suggests that additional items could improve the validity of the DQ, particularly for mild cases. Knowledge of the subject's living situation (living alone, in a nursing home, or with spouse or children), recent driving history, activities of a typical day, and the frequency of contact between the subject and the informant could greatly facilitate early case ascertainment. Further studies with these items are in progress.

Some limitations of this study should be noted. The BLSA is a volunteer cohort with relatively high socioeconomic status and education level. Therefore, informants were able to provide considerable data in a sophisticated fashion. This probably contributed to the high sensitivity of this study. In addition, BLSA subjects provided us with the name and phone number of a suitable informant, facilitating the process considerably. It remains to be seen to what extent these results are generalizable to other populations. Last, the small number of subjects with dementias other than AD limit our ability to comment on the usefulness of the DQ for identification of subjects with these other illnesses. Additional studies in community-dwelling subjects representing a broader range of racial, educational, and medical categories are necessary.

Accepted for publication August 12, 1993.

This study was supported in part by grant RO1 AG08325-01 from the National Institutes of Health, Bethesda, Md, and by the John A. Hartford Foundation, New York, NY.

We thank Florence Kramer for her tireless dedication as an interviewer, William B. Greenough III for his support and inspiration, and the Baltimore Longitudinal Study of Aging participants and scientists who made this work possible.

Reprint requests to Department of Neurology, The Johns Hopkins–Bayview Francis Scott Key Medical Center, 4940 Eastern Ave, Baltimore, MD 21224 (Dr Kawas).

REFERENCES

1. American Psychiatric Association, Committee on Nomenclature and Statistics. *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition.* Washington, DC: American Psychiatric Association; 1987:567.
2. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry.* 1968;114:797-811.
3. 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;12:189-198.
4. Jorm AF, Jacomb PA. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): socio-demographic correlates, reliability, and some norms. *Psychol Med.* 1989;19:1015-1022.
5. Kukull WA, Larson EB. Distinguishing Alzheimer's disease from other dementias: questionnaire responses of close relatives and autopsy results. *J Am Geriatr Soc.* 1989;37:521-527.
6. Davis PB, White H, Price JL, McKeel D, Robins LN. Retrospective postmortem dementia assessment. *Arch Neurol.* 1991;48:613-617.
7. Silverman JM, Breitner JC, Mohs RC, Davis KL. Reliability of the family history method in genetic studies of Alzheimer's disease and related dementias. *Am J Psychiatry.* 1986;143:1279-1282.
8. Silverman JM, Keefe RSE, Mohs RC, Davis KL. A study of the reliability of the family history method in genetic studies of Alzheimer's disease. *Alzheimer Dis Assoc Disord.* 1989;3:218-223.
9. Shock N, Greulich R, Andres R, et al. *Normal Human Aging: The Baltimore Longitudinal Study of Aging.* Bethesda, Md: National Institutes of Health; 1984. Publication 84-250. US Dept of Health and Human Services. *Normal Human Aging: The Baltimore Longitudinal Study of Aging.* Washington, DC: US Dept Health and Human Services; 1984.
10. Grober E, Buschke H. Genuine memory deficits in dementia. *Dev Neuropsychol.* 1987;3:13-16.
11. Kaplan E, Goodglass H, Weintraub S. *The Boston Naming Test.* Boston, Mass: Kaplan & Goodglass; 1978.
12. Benton AL, Hamsher K. *Multilingual Aphasia Examination.* Iowa City: University of Iowa Hospitals; 1976. Revised 1978.
13. Davies A. The influence of age on Trail Making Test performance. *J Clin Psychol.* 1968;24:96-98.
14. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.
15. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol.* 1982;37:323-329.
16. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Resources Task Force on Alzheimer's Disease. *Neurology.* 1984;34:939-944.
17. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1981:321.
18. Lilienfeld AM, Lilienfeld DE. *Foundations of Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1980:375.