

Original Research Article

Testing Episodic Memory in Elderly Subjects: Not as Simple as It Looks

Leo Boelaarts^a Sascha R.A. Meyer^{a, c} Philip Scheltens^b
Jos F.M. de Jonghe^a

^aDepartment of Geriatric Medicine, NoordWest Hospital, Alkmaar, The Netherlands;

^bAlzheimer Center, Amsterdam University Medical Centers, Amsterdam, The Netherlands;

^cSchool of Mental Health and Neuroscience, Maastricht University Medical Center, Maastricht, The Netherlands

Keywords

Aging · Episodic memory · Alzheimer's disease · Neuropsychological tests

Abstract

Background/Aims: The available episodic memory tests are not specifically constructed to examine older subjects. Their use in outpatient memory clinics may result in aborted test administration. We used a strict adherence to the test protocol in cognitively healthy, amnesic mild cognitive impairment (aMCI), and Alzheimer's disease dementia subjects to assess the possibility of preventing this. **Methods:** This is a cross-sectional study in memory outpatient subjects with a mean age of 74.5 years. Primary study outcomes were: number of missing values and test results in the Visual Association Test (VAT) and the 15 Word Test (15WT). **Results:** A strict adherence to the test protocol resulted in a 10-fold decrease in the number of missing values in the VAT. For the 15WT this could not be realized mostly because the test was deemed too demanding for 1 in 6 patients. **Conclusions:** This study is one of the few examining the applicability of well-known episodic memory tests in older subjects. A strict adherence to the test protocol reduced the number of missing values. Floor effects were stronger for the 15WT than for the VAT. Results favor the use of the VAT in senior subjects and show the unsuitability of the 15WT in this group.

© 2019 The Author(s)

Published by S. Karger AG, Basel

Introduction

Elderly subjects presenting with cognitive problems in an outpatient memory clinic are usually assessed using a standard battery of cognitive tests. Results of episodic memory tests may indicate incipient Alzheimer's disease (AD) [1–5].

Few if any episodic memory tests have been specifically developed for elderly subjects. Among the available episodic memory tests, the Rey Auditory Verbal Learning Test (RAVLT) is commonly used [6]. The Dutch adaptation of the RAVLT is the 15 Word Test (15WT) [7], which may be strenuous and easily tiring even for cognitively healthy subjects [8], thereby showing strong floor effects. Age-related changes in cognition are mainly caused by a reduction of the speed of perception and reasoning and especially affect working memory capacity [9]. A measurable decrement in these functions can be found as a linear function of increasing age going from 18 to 50 years. This effect accelerates after the age of 50 years [10, 11].

Meyer et al. [16] demonstrated that the less taxing Visual Association Test (VAT) [12–16] is a sensitive and specific measure of episodic memory function and at the same time shows less floor effects in older subjects compared to the 15WT. Patients who refuse to take a test or who stop the administration of a test due to fatigue because of an advanced age may generate missing test values, potentially leading to result bias.

As far as we know, no studies have reported on the number of elderly patients discontinuing the 15WT or the VAT, or on ways to decrease the number of missing values. We hypothesized that focusing on ways in the administration of the tests to minimize the number of missing values, a strict adherence to the test protocol, is feasible and will also increase test results. In addition, we studied the effect of age on the number of missing values and test results.

Methods

Study Design

This is a cross-sectional study examining the number of missing values in a strict adherence to a test protocol (i.e., the experimental condition) and in routine administration of episodic memory tests (i.e., the control condition).

Patients

Patients were recruited from an outpatient memory clinic population and assessed in a diagnostic program comprised of taking a medical history, a physical examination, a standardized neuropsychological test battery, laboratory tests, ECG, and MRI or CT. Patients were randomly selected to take the VAT or a parallel version of the VAT and took an extra VAT and 15WT subtest; in the present study this comprised the study A group. A smaller control group, study B, was created, comprised of patients who were assessed using the usual test protocol. The patient characteristics and the neuropsychological test battery were described earlier [16, 17]. For the present study, results of patients not able to read letters of 2 mm in height from a comfortable reading distance and patients with a second diagnosis potentially influencing the test results, e.g., depression or anxiety disorder, were excluded.

We used the original VAT version in study A, resulting in roughly half the number of patients in the delayed recall and recognition condition in comparison to the immediate recall condition.

The data used in the present study were of subjects who had received 1 of the following 3 diagnoses: no cognitive disorder (NCD), amnesic mild cognitive impairment (aMCI) according to the Petersen criteria [18], and AD dementia (ADD) using the NINDS-ADRDA criteria [4].

Administration of the VAT and the 15WT

The 15WT is in the test battery used to prevent ceiling effects and the VAT is used to prevent floor effects.

As the goal of study A was to validate the VAT, it was essential to have also data of more severely cognitively impaired patients. Psychotechnicians were specifically instructed in an attempt to minimize the number of missing values but the administration of the tests per se was identical in both protocols and test results were not expected to be different beforehand. In the Appendix, these instructions and the usual instructions for psychotechnicians are shown.

The psychotechnician reported in categories “not administered”, indicating that the patient refused to take the test or the test leader assessing the patient determined that the patient was incapable of taking the test for any reason; “stopped during administration”, denoting refusal of the patient to continue the test; “missing values due to other causes”. This information was not available for study B patients; “missing values” here indicated all-cause missing values.

The Dutch Cognitive Screening Test is administered in our test battery. For this study the results of this test were converted into Mini-Mental State Examination (MMSE) scores. This test and the conversion operation have been well validated [19].

Statistical Analysis

Data were analyzed using the SPSS version 20 statistical package for Windows and the MEDCALC[®] statistical package. Results are expressed as numbers (%) or means ± SD. One-way analysis of variance (ANOVA) was used to study the effect of diagnostic group membership on age and MMSE, as well as the relation between age and membership of a diagnostic group and education level. The χ^2 statistic was used to study the sex ratio and diagnosis. The relation between age and the number of missing values was studied using the Fisher exact test. To compare means of test results the paired samples or independent samples *t* test statistic was used and in case of a small number of cases the nonparametric Kruskal-Wallis test was used. We employed a linear regression analysis to examine the effect of age on the results of the VAT and the 15WT in the 3 diagnostic groups. $p < 0.05$ was considered a statistically significant effect.

Results

Study A (Tables 1, 2)

Clinical Features, Demographics, and Mean Test Results of the Diagnostic Groups

The participating patients may be considered to represent a typical sample of outpatients in a geriatric department; 50% of the patients used more than 3 drugs, and 10–20% of the patients had a history of heart failure, atrial fibrillation, coronary pathology, cerebrovascular and peripheral artery disease, hypertension, hypercholesterolemia, or diabetes.

aMCI and ADD patients were older and had lower MMSE scores (ANOVA, $p = 0.000$ and $p = 0.000$, respectively). Seventy-one percent of the NCD subjects were female, in both of the other diagnostic groups gender was more evenly distributed (χ^2 , $p = 0.018$). Fourteen percent of the subjects had a low education level, 59% had a medium level, and 27% had a higher education. In the univariate analysis of variance, age was a significant factor for diagnostic group membership but education level was not.

Immediate recall as measured with the VAT overtaxed the ability to store and recall information in aMCI and ADD patients as none of these patients scored above the cut-off value of 9. The results in the delayed recall trial were comparable to the immediate recall trial results;

Table 1. Demographics and VAT test results for all patients and diagnosis groups (study A)

	All	NCD	aMCI	ADD
Subjects	182	68	46	68
Age, years	75.5±8.9	70.0±9.1	77.0±7.7	79.9±6.2
Female sex	105±58	48±71	25±54	32±47
MMSE score	24±5.1	28±0.96	25±3.1	20±5.5
<i>VAT immediate recall sum result of 2 trials (maximum score: 12)</i>				
Subjects	178	68	44	66
Not administered	0	0	0	0
Stopped during testing	0	0	0	0
Missing values	4	0	2	2
Mean	7.2±4.7	11.2±1.3	7.3±3.8	2.8±3.4
<i>VAT delayed recall (maximum score: 6)</i>				
Subjects	96	36	30	30
Not administered	0	0	0	0
Stopped during testing	0	0	0	0
Missing values	1	0	1	0
Mean	3.8±2.5	5.8±0.45	3.8±2.1	1.4±2.0
<i>VAT multiple choice recognition (maximum score: 6)</i>				
Subjects	96	36	30	30
Not administered	0	0	0	0
Stopped during testing	0	0	0	0
Missing values	1	0	1	0
Mean	4.8±2.0	6.0±0.17	5.7±0.71	2.5±2.2

Values are presented as numbers or means ± SD.

Table 2. 15WT results for all of the patients and diagnosis groups (study A)

	ALL	NCD	aMCI	ADD
Subjects, n	182	68	46	68
<i>15 WT immediate recall sum result (maximum score: 75)</i>				
Subjects	152	67	46	39
Not administered	29	1	0	28
Stopped during testing	1	0	0	1
Mean	29.3 13.9	40.9 11.1	22.7 7.1	17.3 7.8
<i>15WT delayed recall sum result (maximum score: 15)</i>				
Subjects	151	66	46	39
Not administered	29	1	0	28
Stopped during testing	1	0	0	1
Missing value	1	1	0	0
Mean	4.8 4.7	8.5 3.7	1.9 1.7	1.8 4.5
<i>15WT delayed recognition total result of falsely recognized items (maximum number of errors: 30)</i>				
Subjects	148	66	46	36
Not administered	30	1	0	29
Stopped during testing	3	0	0	3
Missing value	1	1	0	0
Mean faults	5.3 4.7	1.5 2.1	7.5 4.2	9.4 3.3

Table 3. Demographics, VAT, and 15WT test results for all patients and diagnosis groups (study B)

	ALL	NCD	aMCI	ADD
Subjects	45	14	15	16
Age, years	73.5±9.0	65.6±8.3	75.2±6.7	78.8±6.8
Female sex	23±51	4±29	10±67	9±56
MMSE	24±4.5	28±1.5	25±2.5	20±4.1
<i>VAT immediate recall sum result of trials (maximum score: 12)</i>				
Subjects	39	11	15	13
Missing value	6	3	0	3
Mean	7.1±4.1	10.8±2.4	6.3±3.5	4.9±3.8
<i>VAT delayed recall sum result of 2 trials (maximum score: 12)</i>				
Subjects	37	10	14	13
Missing value	8	4	1	3
Mean	4.0±3.0	6.5±2.0	3.8±3.2	2.2±1.9
<i>15WT delayed recall sum result (maximum score: 15)</i>				
Subjects	37	13	13	11
Missing value	8	1	2	5
Mean	3.5±3.6	6.8±4.0	2.2±1.9)	1.2±1.7
<i>15WT delayed recognition total result of falsely recognized items (maximum number of errors: 30)</i>				
Subjects	36	13	13	10
Missing value	9	1	2	6
Mean faults, <i>n</i>	7.1±5.5	3.5±3.7	7.2±3.3	11.9±6.2

Values are presented as numbers or means ± SD.

all aMCI and ADD patients scored below the delayed recall cut-off value of 4. Remarkably, aMCI and ADD patients significantly improved on the VAT recognition trial compared to their delayed recall scores (*t* test, $p = 0.000$). NCD subjects outperformed aMCI patients on 15WT immediate recall, and aMCI patients did better than ADD patients. aMCI and ADD patients did equally poor on 15WT delayed recall. Both groups performed much better on 15WT recognition compared to immediate and delayed recall (paired *t* test, $p = 0.000$).

Missing Values, Effect of Age on Number of Missing Values, and Mean Test Results

None of the patients refused to take the VAT and all of the patients completed this test. Comparing all 3 VAT subtests, we found that maximally 2.2% of all cases had missing values for miscellaneous reasons.

In the 15WT, we found for all cases maximally 18.7% missing values due to all possible causes. In maximally 16.5% of all cases this test was not administered, in all but 1 case this occurred in the ADD group. The test was aborted in 3 ADD subjects, comprising 1.6% of all cases. Missing values were reported due to other reasons in 1 patient, comprising 0.5% of all cases.

Age was not associated with the number of missing values for the VAT. A strong age effect was found for the 15WT number of missing values.

Within each diagnostic group, age was not associated with any of the VAT subtest results. Age was significantly associated with all three 15WT subtest scores in NCD subjects (immediate recall, $p = 0.003$, $R^2 = 12.8\%$; delayed recall, $p = 0.001$, $R^2 = 15\%$; and recognition, $p = 0.001$, $R^2 = 15\%$, respectively). No significant age effects on 15WT scores were found in the aMCI group or in the ADD group.

Study B (Table 3)

Clinical Features, Demographics, and Mean Test Results of the Diagnostic Groups

The education levels and patient characteristics were comparable to those of the study A group; aMCI and ADD patients were older and had lower MMSE scores. In this study 71% of NCD subjects were male, with a higher percentage of women in the other diagnostic groups (χ^2 ns). Mean test results and SD were comparable to those of study A.

Missing Values, Effect of Age on Number of Missing Values, and Mean Test Results

In both VAT conditions in all cases maximally 17.8% missing values were reported. In the NCD group most missing values were reported (maximally 28.6%), in comparison to the aMCI subjects (6.7%) and the ADD group (18.8%).

For the 15WT the maximum percentage of missing values for all cases was 20%. For this test the percentage of missing values was lowest in the NCD group (7.1%) in comparison to aMCI patients (13.3%) and the ADD group (37.5%).

No age effect was found in the number of missing values in either test (χ^2 0.811). The number of patients who had missing values either in the VAT or in the 15WT amounted to 23 (13 younger and 10 older than the mean age of 73.5 years).

Within all diagnostic groups no age effect was found on any mean test result, with one exception, i.e. the control study B group in the 15WT delayed recognition condition ($p = 0.037$, $R^2 = 34\%$).

Comparison of Mean Test Results within Diagnostic Groups between Study A and Study B

Data for comparison were available for the VAT immediate recall and 15WT delayed recall and recognition conditions. The only significant differences were found in ADD patients who scored significantly higher in study B in the VAT immediate recall (t test, $p = 0.03$), and in the 15WT delayed recognition condition the control group in study A scored significantly better than in study B (t test, $p = 0.008$).

Discussion

In 2 studies in healthy, aMCI, and ADD subjects we found a profound effect of strict adherence to the protocol on the number of missing values. Age did not affect the number of missing values or the mean test results, with the exception of younger subjects (they performed best in the 15WT).

For the VAT the all-cause number of missing values in study A was maximally 2.2. and it was 17.8% in study B. Remarkably, the highest number of missing values in study B was found in NCD subjects, which could be caused by a ceiling effect, e.g., growing bored or experiencing the test as childish. This effect could, however, be cancelled out in this group by the study A protocol. In the study B ADD group, even in the less demanding VAT, we found a high number of 18.8% of missing values. In study A this was maximally 3.0%.

For the 15WT the all-cause numbers of missing values were similar in both studies. This was caused by the high percentage of missing values in ADD subjects. Given the low percentage of ADD patients who stopped the test during administration, the high percentage of “not administered” annotations in study A represented the test leader’s decision not to start the test. This may illustrate the unsuitability of this test for use in ADD subjects. The success of the study A test protocol is also illustrated by the low number of missing values in the NCD and aMCI subjects in comparison to study B.

Based on our findings we conclude that the number of missing values for test results of senior NCD, aMCI, and ADD subjects may be reduced if tests are not too strenuous, burdening their limited memory capacity too much [8–11]. These results also highlight the importance of specifically training psychotechnicians to test this group of senior patients.

The second conclusion is that strict adherence to the test protocol did not generally improve test results. The apparent contradiction of the significantly better VAT immediate recall test results for the ADD group in study B may well be caused by the success of the study A test protocol. In study A also more demented patients completed their tests, which probably led to lower mean test results for this group in comparison to the ADD study B group. In accordance with this notion, the controls in study A performed better in the 15WT delayed recognition subtest; a strict adherence to the test protocol can help subjects to achieve better.

At the individual patient level, the effect of the attempts to decrease the number of missing values will be less important than the profile of test results and test behavior. However, when categorizing subjects into diagnostic groups based on test results or examining efficacy of new therapies in different diagnostic groups, the number of missing values and better test results may be critically important.

The third conclusion is that age did not affect the number of missing values; for the VAT this was probably due to the small number of missing values. For the 15WT this result may be confounded by the older, ADD subjects. Importantly we did not compare subjects under 50 years of age or over 80 years; the mean age was 73.5 years, meaning that all of the subjects may be characterized as senior.

The fourth conclusion is that we did not find an age effect on the mean test results, with the exception of the group of controls. The VAT is designed to minimize floor effects and thereby the effect of age. By contrast, in accordance with the literature, younger subjects in this study performed better on the 15WT because this test draws heavily on working memory capacity and speed of processing [8]. Also the effect of the diagnosis of aMCI or ADD probably masks any age effect with this number of subjects.

Most of the available episodic memory tests such as the RAVLT have been developed with the aim of assessing younger subjects suffering of varying ailments. Spaan et al. [20] provided an overview of the available episodic memory tests and concluded that episodic memory cannot be adequately tested when a neurodegenerative disorder is suspected due to the fact that these tests were not developed to just measure this. Cerami et al. [21] examined the extent to which episodic memory tests can be used as biomarkers. Rigorous criteria used to evaluate the methodology in pharmaceutical research were applied to evaluate the robustness of episodic memory tests in this respect. All tests failed to fulfil these criteria.

The aim of a test battery is to obtain reliable data in a standardized way, which means that repeated testing will yield identical results. This advantage is balanced by a decrement in ecological validity. An extensive assessment of cognitive function using demanding tests may produce a distorted representation of the actual cognitive ability. Even worse, submitting senior patients to demanding tests may discourage them, leading to interference with the participation observation by the psychotechnician and loss, in this way, of important information to make a clinical assessment of behaviour and cognition.

If we were to generalize the 15WT results of this study, it could mean that, even before starting the test administration in the usual way in aMCI and ADD senior subjects, a priori chances are that in 15–45% of the cases no test results will be generated. Even when using a strict adherence to the test protocol the a priori chance of missing out on 15WT test results may still be around 40% in ADD subjects. This underlines the unsuitability of this test to study episodic memory especially in these patients.

Notably, we found that in ADD and aMCI patients recognition is much better than free recall. Perhaps this is due to a milder stage of AD in our patients causing a less compromised

learning ability compared to patients described in the older literature who already suffered from a more advanced stage of the disease.

aMCI and ADD patients showed comparable results in the VAT and delayed 15WT recall condition. Previous research has shown an improved memory performance by these subjects when retrieval support by means of cueing or recognition is provided [22]. This illustrates a continuum of episodic memory test performance that is mildly to severely impaired, going from the aMCI to the ADD stage [23].

The uneven gender distribution across the diagnostic groups, especially in the control group, is coincidental as all patients were selected from the same outpatient clinic. Normative data show that women perform somewhat better on verbal learning tests such as the 15WT [8]. However, the diagnosis will have a more profound effect on test results and will probably dwarf these gender effects.

To summarize, the main conclusions of this study are the positive effect of a strict adherence to the test protocol and the unsuitability of the 15WT to test episodic memory in elderly aMCI and ADD subjects.

Our results underline the need to develop more ecologically valid tests for this group. In the meantime, we would like to propose the use of the VAT as an episodic memory test in these subjects.

Statement of Ethics

The study protocol was approved by the local ethics committee. All of the patients gave written informed consent.

Disclosure Statement

S. Meyer and J. de Jonghe are the authors of the parallel and extended versions of the VAT.

Funding Sources

No funding was received.

Author Contributions

L. Boelaarts designed this study, analyzed the data, and wrote this article. S. Meyer carried out this study, built the data base and collected the data, and assisted in writing this article. Ph. Scheltens assisted in designing this study and in writing this article. J. de Jonghe assisted in designing this study, analyzing the data, and writing this article.

Appendix

Instructions for Psychotechnicians

Routine clinical cognitive testing requires that psychotechnicians use standard test instructions. They are instructed to use a standard test protocol and get objective information on multiple cognitive domains. If a patient finds it difficult to understand test instructions or if he or she fails during the try-out phase, this test will not be completed.

Test results will be discussed by the psychotechnicians and the senior neuropsychologist in multidisciplinary consensus meetings and they will be presented by the physician in the visit when all of the results of the diagnostic program are presented to the patient and informant.

Enhanced Instruction

The awareness of the psychotechnicians regarding the importance of the full completion of the VAT and the 15WT was raised by the research leader of study A by explaining the study goals of study A but also the potential benefit for the patient completing all of the tests and thereby gaining a reliable assessment of the cognitive functions of the subject.

Subsequently the psychotechnicians were trained to extra motivate and support patients. It was explained that the assessment of the psychotechnician of the ability of a patient to take a test remains subjective and that patients may very well be capable of taking a test or achieving better results if they are properly supported and encouraged to give it a try and do the best they can.

Instructions to the patient at the start of the test administration could be as follows: “it is important to give your best to better understand your complaints and the possible causes for these problems.”

The psychotechnicians were instructed to focus on complimenting patients during administration for correct answers and for their effort in achieving as good as possible results, for example, by saying: “that answer is correct,” and “well done for giving your best.” Also, when experiencing a loss of focus, patients were instructed as follows: “please try again,” “well done, this answer is indeed correct,” “well done, this answer is also correct,” and “take your time, no need to rush.”

Each session was evaluated afterward by the psychotechnician and the research leader to potentially minimize further missing values and maximize test achievement.

References

- 1 Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer’s disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol*. 2007 Aug;6(8):734–46.
- 2 Hort J, O’Brien JT, Gainotti G, Pirttila T, Popescu BO, Rektorova I, et al.; EFNS Scientist Panel on Dementia. EFNS guidelines for the diagnosis and management of Alzheimer’s disease. *Eur J Neurol*. 2010 Oct;17(10):1236–48.
- 3 Knopman DS, DeKosky ST, Cummings JL, Chui H, Corey-Bloom J, Relkin N, et al. Practice parameter: diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*. 2001 May;56(9):1143–53.
- 4 McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. *Alzheimers Dement*. 2011 May;7(3):263–9.
- 5 Sorbi S, Hort J, Erkinjuntti T, Fladby T, Gainotti G, Gurvit H, et al.; EFNS Scientist Panel on Dementia and Cognitive Neurology. EFNS-ENS Guidelines on the diagnosis and management of disorders associated with dementia. *Eur J Neurol*. 2012 Sep;19(9):1159–79.
- 6 Rey A. The clinical examination in psychology. Paris: Presses Universitaires de France; 1958.
- 7 Saan RJ, Deelman B. 15-Woorden Test A en B [15-Word test A and B.] Bouma A, Mulder J, Lindeboom J, Schmand B, editors. *Handboek neuropsychologische diagnostiek* [Handbook neuropsychological assessment.] 2nd ed. Amsterdam: Pearson Assessment and Information BV; 1986. pp. 267–82.
- 8 Van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J. Rey’s verbal learning test: normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *J Int Neuropsychol Soc*. 2005 May;11(3):290–302.
- 9 Bors DA, Forrin B. Age, speed of information processing, recall, and fluid intelligence. *Intelligence*. 1995;20(3):229–48.
- 10 Verhaeghen P, Salthouse TA. Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychol Bull*. 1997 Nov;122(3):231–49.
- 11 Bopp KL, Verhaeghen P. Aging and verbal memory span: a meta-analysis. *J Gerontol B Psychol Sci Soc Sci*. 2005 Sep;60(5):223–33.

- 12 Lindeboom J, Schmand B. [Visual association test: Manual](#). Leiden: PITS; 2003.
- 13 Lindeboom J, Schmand B, Meyer S, de Jonghe JF. [Visual association test: Manual](#). Amsterdam: Hogrefe; 2014.
- 14 Lindeboom J, Schmand B, Tulner L, Walstra G, Jonker C. Visual association test to detect early dementia of the Alzheimer type. [J Neurol Neurosurg Psychiatry](#). 2002 Aug;73(2):126–33.
- 15 Meyer SR, de Jonghe JF, Schmand B, Ponds RW. The Visual Association Test-Extended: a cross-sectional study of the performance validity measures. [Clin Neuropsychol](#). 2017 May;31(4):798–813.
- 16 Meyer SR, Spaan PE, Boelaarts L, Ponds RW, Schmand B, de Jonghe JF. Visual associations cued recall: a paradigm for measuring episodic memory decline in Alzheimer’s disease. [Neuropsychol Dev Cogn B Aging Neuropsychol Cogn](#). 2016 Sep;23(5):566–77.
- 17 Boelaarts L, Scheltens P, de Jonghe J. Does MRI increase the diagnostic confidence of physicians in an outpatient memory clinic [doi]. [Dement Geriatr Cogn Disord Extra](#). 2016 Jun;6(2):242–51.
- 18 Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. [Arch Neurol](#). 1999 Mar;56(3):303–8.
- 19 Schmand B, Deelman B, Hooijer C, Jonker C, Lindeboom J. The items of the cognitive screening test compared to those of the Mini-Mental State Examination. [Tijdschr Gerontol Geriatr](#). 1996;29:29–33.
- 20 Spaan PE, Raaijmakers JG, Jonker C. Alzheimer’s disease versus normal ageing: a review of the efficiency of clinical and experimental memory measures. [J Clin Exp Neuropsychol](#). 2003 Apr;25(2):216–33.
- 21 Cerami C, Dubois B, Boccardi M, Monsch AU, Demonet JF, Cappa SF; Geneva Task Force for the Roadmap of Alzheimer’s Biomarkers. Clinical validity of delayed recall tests as a gateway biomarker for Alzheimer’s disease in the context of a structured 5-phase development framework. [Neurobiol Aging](#). 2017 Apr;52:153–66.
- 22 Baeckman L, Jones S, Berger A, Jonsson Laukka E, Small J. Cognitive impairment in preclinical Alzheimer’s disease: a meta-analysis. [Neuropsychol](#). 2015;19(4):520–31.
- 23 Greenaway MC, Lacritz LH, Binegar D, Weiner MF, Lipton A, Munro Cullum C. Patterns of verbal memory performance in mild cognitive impairment, Alzheimer disease, and normal aging. [Cogn Behav Neurol](#). 2006 Jun;19(2):79–84.