Published in: Journal of neurology (2005), vol. 252, iss. 1, pp. 47-55
Status: Postprint (Author's version)

Memory evaluation with a new cued recall test in patients with mild cognitive impairment and Alzheimer's disease

Adrian Ivanoiu Stephane Adam Martial Van der Linden Eric Salmon Anne-Claude Juillerat Reinhild Mulligan Xavier Seron

Abstract

Free delayed recall is considered the memory measure with the greatest sensitivity for the early diagnosis of dementia. However, its specificity for dementia could be lower, as deficits other than those of pure memory might account for poor performance in this difficult and effortful task. Cued recall is supposed to allow a better distinction between poor memory due to concurrent factors and impairments related to the neurodegenerative process. The available cued recall tests suffer from a ceiling effect. This is a prospective, longitudinal study aiming to assess the utility of a new memory test based on cued recall that avoids the ceiling effect in the early diagnosis of Alzheimer's disease (AD). Twenty-five patients with mild cognitive impairment (MCI), 22 probable AD patients (NINCDS-ADRDA) at a mild stage, 22 elderly patients with subjective memory complaints (SMC) and 38 normal age-matched controls took part in the study. The patients underwent a thorough cognitive evaluation and the recommended screening procedure for the diagnosis of dementia. All patients were re-examined 12-18 months later. A newly devised delayed cued recall test using semantic cues (The RI48 Test) was compared with three established memory tests: the Ten Word-List Recall from CERAD, the "Doors" and the "Shapes" Tests from "The Doors and People Test Battery". Forty-four % of the MCI patients fulfilled criteria for probable AD at follow-up. The RI48 Test classified correctly 88 % of the MCI and SMC participants and was the best predictor of the status of MCI and mild AD as well as the outcome of the MCI patients. Poor visual memory was the second best predictor of those MCI patients who evolved to AD. A cued recall test which avoids the ceiling effect is at least as good as the delayed free recall tests in the early detection of AD.

Key words: Alzheimer's disease · mild cognitive impairment · memory

INTRODUCTION

The best cognitive predictor of the development of future dementia is a low memory performance [7, 31]. It has been shown that the tests based on the free recall of newly acquired information are sensitive for the detection of incipient AD [34]. However, these tests could be influenced by concurrent factors, such as advanced age, anxiety, depression, general illnesses, and medications. All those factors may lead to impaired attention, poor strategic search in memory and reduced processing capacity, ultimately resulting in poor performance on the free recall for non-mnesic reasons. The cued recall technique [5, 13] aims at enhancing the spontaneous free recall by the presentation of the same semantic cues to help both the encoding and the retrieval phase of the test. This technique is supposed to minimize the effect of impaired attention, inefficient strategies, or reduced processing capacity and to disclose "true" memory deficits, which are eventually related to incipient dementia. Whereas the impairment on the free recall may not be very specific for dementia, as it can be observed in other conditions, such as aging and depression, the cued recall impairment may appear as a more specific "marker" of the presence of early dementia of AD type. The first test created following the cued recall technique (FCSRT) [5, 13] proved to be reliable and useful in the diagnosis of established AD [8, 19, 26, 33].However, its sensitivity to minimal impairment is limited by the presence of a ceiling effect in controls and, presumably, in some highly functioning incipient AD patients [18]. Subsequently, a test using 64 items instead of 16 as in the FCSRT showed good validity in the diagnosis of AD,bet-ter than classical memory tests [4]. However, some limitations of this study must be stressed, such as the heterogeneous sample, the absence of any follow-up, a limited number of poorly characterized "very mild" AD patients and a procedure that is cumbersome for clinical application. In order to avoid ceiling effects, while at the same time having a simpler test, feasible in clinical settings,we devised a 48 items delayed cued recall test (The RI48 Test), along the same line of that proposed by Buschke [1, 4]. The goal of the present study was to evaluate the diagnostic validity for "mild cognitive impairment" (MCI) [27] and incipient AD of the RI48 Test, when compared with other established and frequently used memory tests. It was predicted that the RI48 Test might be a good diagnostic test for MCI and incipient AD, as it might allow a better separation between AD and the control group.

SUBJECTS AND METHODS

The participants were selected amongst patients attending one of the three Memory Clinics taking part in the study (Brussels, Liège, and Geneva).All patients had French as their first language and all centers used the same

procedure. Those patients who were aged > 80, or had severe dementia, non-AD dementia, neurological or psychiatric conditions without dementia, severe general illnesses or incomplete data were excluded from the study. Eighty-five patients were included.The initial screening included a clinical examination by a trained clinician, a detailed interview with an informant and cognitive screening using the Mini Mental State Examination (MMSE) [10] and the Mat-tis Dementia Rating Scale (DRS) [22], adapted in French. The screening procedure for dementia followed published guidelines [28]. The Clinical Dementia Rating Scale (CDR) [17] was used to assess dementia severity. A brain morphological imaging scan (CT or MRI) and a complete neuropsychological examination were carried out in each case. The participants were classified in three groups, according to their diagnosis: probable AD(PrAD) [23], mild cognitive impair-ment(MCI) [27] and patients with subjective memory complaints (SMC). The criteria for MCI [27] were required. For the diagnosis of MCI the performance on the memory evaluation had to be less than the cut-off corresponding to –1.5 SD for either the subtest "recall" of the MMSE (cut-off < 2/3) or the subtest "memory" of the DRS (cutoff < 22/25). Those patients complaining of memory loss but without objective impairment received the diagnostic label of "subjective memory complaints" (SMC), as they did not fulfil criteria for AD or MCI. Besides the absence of dementia, the SMC patients were diagnosed following two criteria: the presence of an isolated subjective memory complaint in the absence of an objective memory impairment, that is a score at both screening memory tests (MMSE and the DRS memory subtests) above or equal to –1.5 SD when compared with healthy elderly persons. The presence of significant non-mnesic symptoms was an exclusion criterion. All individuals in this latter group were free of major psychiatric or neurological illnesses but all of them expressed some anxious feelings about senescence, age related cognitive decline or Alzheimer's disease. Sixty-nine patients out of 85 (81%) completed the study by returning for follow-up evaluation after 12 to 18 months (mean ± SD: 14.9 ± 2.6). All PrAD patients were confirmed as PrAD and all individuals remaining in the SMC group were stable, within the normal range.Amongst the 25 MCI patients, 11 (44%) deteriorated, fulfilling criteria for probable AD. The remaining 14 (56%) MCI patients showed no clear impairment on tests or a significant decline in their day by day functioning, although relatives often reported some subtle impairment. None of them im-proved.As these MCI patients did not fulfill criteria for AD at follow-up, they were considered as "stable" MCI.We performed a first analysis on the three groups of patients (PrAD, MCI and SMC), compared with a group of 38 normal elderly (NE), community dwelling volunteers. Subsequently, we carried on further analysis with two additional groups:

– the "confirmed" AD group (cfAD), formed by pooling together the MCI patients who deteriorated to AD at follow-up and the PrAD. We considered those MCI who evolved to AD to have, in fact, very mild AD pathology at the moment when they were first assessed. This classification allowed us to take into account the split of the MCI group into evolving and stable MCI, without the inconvenience of handling groups of very small size. – the "potential" AD group (ptAD), formed by all the MCI and PrAD patients. We considered this last group because the MCI patients, as defined in this study, supposedly have underlying AD pathology for a majority of them. Evidence from previous studies showed that as many as 75–80% of the MCI patients diagnosed according to the same criteria as ours are likely to evolve to AD in five years time [27].As MCI and AD are clinical concepts, whereas the passage from normal cognition to dementia is presumably a continuum rather than sharply demarcated, we assumed it is appropriate to pool together the pathological groups. – The memory evaluation included the "Rappel Indicé 48 items" (RI48), which was devised by Adam et al. [1] as a simplified and shortened form of the test developed by Buschke [4]. The task comprised 48 different items, belonging to 12 different semantic categories (four words for each of the 12 categories; e.g.., the "weapon" category had the words for "crossbow", "dagger", "blud-geon", and "pistol"). The 48 items were presented to participants as written words on 12 consecutive cards, each card containing four items, each item from a different category (e. g. the first card contained the French words for an insect - "ladybird", a fruit -"raspberry", a tree -"palm" and a garment -"jacket". Participants were asked to encode these items with the help of semantic cues: for example, the word "dentist" will be encoded in relation with the category "profession". On completion of each board, an immediate cued recall test is performed. After the last board is shown, participants are asked to count backward for 20 seconds. On completion of this latter task, participants perform a cued recall task, using the categories as cues, e. g. "which were the flowers, insects, etc.". The RI48 Test took 20 to 25 minutes to administer, which is exactly the same time as the available cued recall test, the FCSRT [8, 19]. Performance on the RI48 Test was compared with that on several well-known episodic memory (EM) tests, widely used in previous studies with AD patients. The choice of EM tests and scoring procedures was made so as to have different types of item (verbal or visual) and different encoding and retrieval modes (free, cued, recall, recognition). The first test used is a French adaptation by our team of the Ten Word List Learning and Recall from the CERAD battery [25]. Three scores resulted from this procedure: the total number of words recalled in three immediate recalls (immediate recall - IR), the number of words produced in the delayed recall phase (delayed recall - DR) and the difference between the delayed recall and the score of the last trial ("savings"). The two other tests used were two subtests evaluating visual memory from the "Doors and People" Battery [3]. The "Shapes" Test evaluates the recall of four simple figures. In the same manner as for the CERAD test, three types of score were used: the sum of the three immediate recalls of figures (IR), the delayed recall (DR), and the difference between the delayed

recall and the score on the last trial of the test ("the savings"). The second test is a nonverbal recognition task involving the presentation of two series of color photographs of doors, which subsequently had to be pointed out from four alternatives. The first set of 12 doors (set A) is easier than the second one (set B). A full neuropsychological evaluation was performed on the first visit by a trained neuropsychologist who was unaware of the clinical diagnosis. The evaluation of EM was considered separately and was not used in the diagnostic decision process. The neuropsychological assessment included tests evaluating: language and semantic memory (The Category Fluency Test for animals, the LEXIS Naming Test) [6], visuo-spatial processing (the "Clock Drawing Test", The "Praxis" part of the CERAD battery) [25, 30], attention (The "d2" cancellation Test) [16] and executive functions (The Letter Fluency Test for the letter P [6], the "Stroop" Test and the "Trail Making Test" - these last two tests were adaptations in French by our team, unpublished). All these cognitive tests are classically used in the assessment of AD patients [21, 25].

STATISTICS

A one-way ANOVA, with each cognitive measure as the dependent variable and a factor representing the four groups, was performed using Stat View [32]. The relevant pairwise comparisons (Bonferroni-Dunn) were carried out between groups (NE vs SMC, SMC vs. MCI, SMC vs. PrAD, MCI vs. PrAD). The diagnostic sensitivity/specificity, overall hit rate (percentage of subjects correctly classified by the test) as well as the positive and negative predictive values [11] of each test were computed with the z-scores method [24]. The z-score represents, in standard deviation units, the amount a score deviates from the mean of the control population: z-score = (subject's score - controls' mean)/controls' SD. The performance was considered as impaired at a z-score of less than -1.96. When the specificity was not at 100%, the likelihood ratios (LR) were computed with the formula: LR = sensitivity/(1-specificity). For three selected memory tests, supposedly the most sensitive ones, the Receiver Operating Characteristic (ROC) curve was also performed [15, 29]. Finally, the clinical status was predicted by using a logistic step by step regression (F for enter = 4; F for exit = 3.996; ascendant method) [32].

RESULTS

The demographic data of the patients who completed the study showed no significant differences in terms of age, sex or education (Table 1). We proceeded to the analysis following two steps. First, we compared the four clinical groups, as defined at baseline (NE, SMC, MCI and PrAD). Second, we took into account the follow up data, including in the analysis the MCI subgroups (stable and deteriorated) as well as the pooled groups of cfAD and ptAD. The two steps of the analysis are detailed below.

INTERGROUP COMPARISON

The NE and SMC groups were similar in their performance on all cognitive tests (Table 1). The MCI patients had impaired performance on both dementia screening scales, when compared with both the NE and the SMC group (Table 1). This finding was not surprising, since subtests of these scales were used in the classification of these individuals as MCI. The mean Mini Mental State score of the MCI group was 26.7/30 with a minimum of 24, suggesting that cognition was very slightly impaired in these individuals. The only significant difference observed between MCI and both the NE and SMC groups was on the Category Fluency Test (names of animals evoked freely/2 min), which is supposed to assess the integrity of the "semantic system". On the Naming test, the MCI patients showed a poorer performance than that of NE but not that of SMC. The performance on the memory tests is displayed in Table 2. No significant difference was observed between NE and SMC for any of these measures. The MCI patients were significantly impaired on all verbal memory measures, but their performance was within normal range for the savings of the "Shapes" Test and for the "Doors" Test. The PrAD group was impaired on all memory measures. The sensitivity/specificity, the overall hit rate as well as the positive and negative predictive values of the four different memory tests for the diagnosis of either MCI or PrAD, when compared with the SMC, are shown in Table 3. When the PrAD were taken into account, the RI48 Test, the CERAD-IR and DR as well as the "Shapes"-IR and DR were very close in their ability to separate the AD patients from the SMC group. However, the RI48 Test showed the best overall hit rate, meaning the largest number of participants correctly classified as to their diagnosis (98%). The "savings" measures from both the verbal (CERAD) and the visual ("Shapes") tests were clearly less sensitive to diagnose PrAD. The same observation holds for the "Doors" Test, even when the most difficult part of it was taken into account. Concerning the MCI group, the tests showing the best sensitivity were those verbal memory tests involving delayed recall: CERAD-DR and the RI48 Test. However, the RI48 Test showed a better overall hit rate than the CERAD-DR (88 % of participants correctly classified by the first compared to 83 % for the second). The last finding results from a lower specificity for the CERAD-DR when compared with the RI48, as some SMC subjects are found to be impaired on the first one but not on the second. Visual memory tests were less sensitive for the diagnosis of MCI than the verbal ones. The savings scores from either the verbal or the visual memory tests were not sensitive enough for the diagnosis of MCI. We also performed a similar analysis for the MCI and

the PrAD when compared with a group formed by the SMC and NE patients pooled together. The resulting sensitivity/specificity, overall hit rate as well as the positive and negative predictive values were very similar to those already mentioned and favored the RI48 test. This later analysis allowed us to compute the likelihood ratios, not available for all tests from the comparison with the SMC group alone because for some tests, including the RI48, the specificity values were at 100% (see Table 3). In the case of the MCI patients, the highest LR observed was that of the RI48 Test (LR = 41.5), followed by the CERAD-IR (LR= 19.7). The CERAD-DR came only in the fourth position (LR = 11.9), essentially because of its lower specificity. The same situation was apparent for the PrAD group: LR = 51.0 for the RI48,26.4 for the CERAD-IR test and 14.8 for the CERAD-DR.

Table 1 *General characteristics and the cognitive evaluation of patients and controls*

	Normal Elderly	Subjective Memory Complaints	Mild Cognitive Impairment	Probable AD
Number	38	22	25	22
Age/years	70.8 (8.4)	68.8 (6.1)	71.0 (6.0)	72.9 (4.2)
Sex (% F)	58	64	55	75
Education % 1/2/3[1]	38/33/29	32/36/32	24/45/31	37/42/21
Screening data				
MMSE (0–30)	28.1 (1.8)	28.6 (1.3)	26.7 (1.7)[a, b]	23.0 (2.6)[a, b, c]
DRS (0–144)	138.6 (2.5)	138.7 (3.9)	130.9 (5.9)[a, b]	120.2 (8.7)[a, b, c]
DRS memory (0–25)	24.1 (0.8)	23.8 (1.3)	19.5 (3.5)[a, b]	16.2 (3.0)[a, b, c]
CDR 0/0.5/1/2	–	20/2/0/0	0/25/0/0	0/5/14/3
CDR sum of boxes[2]	–	0.1 (0.2)	1.7 (0.8)[b]	4.3 (2.2)[b, c]
Neuropsychological Tests				
Animals Fluency/2min	29.9 (8.3)	29.7 (7.6)	22.5 (6.6)[a, b]	14.1 (5.6)[a, b, c]
Letter fluency (P)/2 min	20.2 (5.9)	21.5 (6.2)	18.0 (6.2)	14.2 (6.1)[a, b]
Naming Test/64	57.1 (3.9)	56.3 (3.7)	53.8 (4.7)[a]	46.6 (7.3)[a, b, c]
"d2" Test (nb. correct)	332.7 (91.8)	344.9 (86.2)	307.2 (98.2)	263.3 (99.8)
Stroop Test[3]	66.7 (30.4)	68.9 (31.1)	80.7 (45.4)	149.2 (75.4)[a, b, c]
Trail Making Test (B-A)[4]	69.4 (33.7)	65.4 (46.5)	91.9 (51.8)	127.6 (76.1)[a, b]
Clock Drawing Test/10	9.1 (1.5)	9.3 (0.9)	8.4 (1.9)	7.0 (2.3)[a, b, c]
CERAD Figures/11	10.3 (1.2)	10.4 (0.8)	9.8 (1.4)	9.3 (1.7)

The values are mean(SD); [1] 1 = primary school/2 = secondary school/3 = more than secondary school; [2] dementia severity was expressed by two indexes: an algorithm derived from the global CDR score (0/0.5/1/2) and the sum of the different subscores, also known as the "sum of boxes", which is supposed to be more sensitive; [3] the interference score is displayed; [4] the difference between the time necessary to perform part B (tracking by alternating number and letter) and that of part A (simple tracking). One-way ANOVA, with each cognitive measure as the dependent variable and a factor representing the four groups was performed. The relevant pairwise comparisons (Bonferroni-Dunn) were carried out between adjacent groups (NE vs SMC, SMC vs MCI, MCI vs PrAD). The presence of a letter as an exponent indicates a statistical significant difference (at $p < 0.05$) between that performance and the performance of: [a] normal aged; [b] subjective complainers; [c] MCI. Only the meaningful comparisons are shown

Published in: Journal of neurology (2005), vol. 252, iss. 1, pp. 47-55
Status: Postprint (Author's version)

Table 2 *The performance of patients and controls on memory tests*

	Normal Elderly	Subjective Memory Complaints	Mild Cognitive Impairment	Probable AD
RI 48 Test				
• delayed recall/48	25.8 (4.5)	25.6 (4.5)	14.1 (4.7)[a, b]	8.7 (5.4)[a, b, c]
CERAD Word List Recall				
• sum of trials 1-3/30	21.7 (3.3)	21.0 (2.7)	14.7 (3.2)[a, b]	10.5 (3.1)[a, b, c]
• delayed recall/10	7.0 (1.5)	6.7 (1.9)	2.8 (2.3)[a, b]	0.9 (1.3)[a, b, c]
• savings (%)	86.1 (17.5)	81.9 (19.4)	44.0 (29.9)[a, b]	21.9 (30.7)[a, b, c]
"Shapes" Test				
• sum of trials 1-3/36	31.9 (4.5)	31.5 (5.3)	21.9 (9.7)[a, b]	11.5 (6.0)[a, b, c]
• delayed recall/12	11.5 (1.5)	11.1 (1.6)	7.0 (3.7)[a, b]	3.2 (2.7)[a, b, c]
• savings (%)	99.4 (6.0)	99.2 (7.9)	79.3 (20.1)	59.5 (35.2)[a, b, c]
"Doors" Test				
• part A/12	10.3 (1.6)	10.4 (1.4)	8.7 (3.0)	7.3 (2.5)[a, b]
• part B/12*	7.0 (1.8)	7.0 (2.1)	5.2 (2.8)	3.5 (1.7)[a, b]

The values are mean(SD); * Part B of the test was only administered to those patients performing at > 6/12 on part A. One-way ANOVA, with each cognitive measure as the dependent variable and a factor representing the four groups was performed. The relevant pairwise comparisons (Bonferonni-Dunn) were carried out between adjacent groups (NE vs SMC, SMC vs MCI, MCI vs PrAD). The presence of a letter as an exponent t indicates a statistical significant difference (at $p < 0.05$) between that performance and the performance of: [a] normal aged; [b] subjective complainers; [c] MCI. Only the meaningful comparisons are shown

Table 3 *Specificity, sensitivity, overall hit rate, positive and negative predictive values of four different memory tests in the early diagnostic of AD*

		Mild cognitive impairment				Probable AD			
	Spec.	Sens.	OHR	PPV	NPV	Sens.	OHR	PPV	NPV
MMSE	100	12	53	100	50	73	87	100	79
Rl 48 Test									
• delayed recall	100	77	88	100	81	94	98	100	96
CERAD									
Word List Recall									
• sum of trials	100	67	83	100	75	89	95	100	91
• delayed recall	86	81	83	85	82	100	93	86	100
• savings	90	62	76	87	70	79	85	88	83
"Shapes" Test									
• sum of trials	94	54	69	94	55	94	94	94	94
• delayed recall	94	43	61	92	48	94	94	94	94
• savings	100	22	51	100	43	44	71	100	62
"Doors" Test									
• part A	86	46	64	80	56	71	79	83	75
• part A*	86	33	60	70	56	50	73	67	75
• part B	86	42	65	73	62	64	78	70	82

The MCI and PrAD patients were compared to the SMC group. The values are %; *Spec,* specificity; *Sens,* sensitivity; *OHR* Overall Hit Rate (percentage of subjects correctly classified by the test); *PPV and N*Pl/Positive and Negative Predictive Values; The specificity values are identical for the comparison MCI vs. SMC and PrAD vs SMC because the reference population is the same. * those subjects who performed both A and B parts, in order to allow a direct comparison of parts A and B (part B of the test was only administered to those patients performing at > 6/12 on part A

Published in: Journal of neurology (2005), vol. 252, iss. 1, pp. 47-55
Status: Postprint (Author's version)

Additional analysis by taking into account the MCI subgroups

At follow-up, some MCI patients evolved to probable AD whereas others remained stable. Their characteristics are shown in Table 4. Those with MCI who evolved to AD and those who remained stable were not different for age, education or sex. The MMSE and DRS did not differentiate between deteriorating versus stable MCI. Those with MCI who evolved to AD had low initial visual memory, significantly impaired when compared with controls, whereas their verbal memory was virtually indistinguishable from that of PrAD patients. In order to assess the best sensitivity/specificity balance for those memory tests showing the most promising diagnostic qualities we carried out a separate analysis for the RI48 test together with the IR and DR of CERAD, taking the MMSE score as witness and using the ROC curve method (Fig. 1). As we aimed to disclose the best diagnostic potential for the AD pathology since its first clinical manifestations, we contrasted in this analysis the ptAD and the SMC groups. Visual inspection of the graph shows that the RI48 Test has the best balance between sensitivity and specificity for any of the possible cut-off points and the larger area under the curve (AUC). Given the small difference between tests, however, the statistical power allowed by our sample size was not sufficient to confidently reject the null hypothesis. In order to determine if the RI48 Test was able to predict the status of AD better that the other tests a step by step logistic regression analysis was carried out.The six main memory measures were entered as independent variables (RI48, IR and DR of CERAD, IR and DR of the "Shapes" Test and the "Doors" Test part A). The clinical status, as mirrored by the groups and subgroups of patients, was the dependent variable. The following situations were considered and analysed separately: – PrAD vs. SMC (adjusted R^2 =0.87). The status of established AD at a mild stage was best predicted by the IR of the "Shapes" Test (F exit = 18) as well as the IR from the CERAD battery (F exit = 5.5). – cfAD vs. SMC (adjusted R^2 = 0.85). The status of AD when the very mild cases were also included was best predicted by the RI48 Test (F exit = 82). The second predictor was a visual memory test, the "Doors" Test (F exit = 20). – ptAD vs. SMC (adjusted R^2 =0.71).Assuming that the stable MCI were in fact incipient AD,AD and MCI patients were considered in a common group. The RI48 Te s t r e s u l t e d as the sole predictor of the AD "status", as reflected by the ptAD group (F exit = 92).
– MCI vs. SMC (adjusted R^2 = 0.73). The status of MCI, when considered as a separate group, was best predicted only by the RI48 Test (F exit = 70).
– "stable" MCI vs. SMC (adjusted R^2 = 0.68). The status of stable MCI was best predicted only by the RI48 Test (F exit = 41).
– "stable"MCI vs.PrAD (adjusted R^2 = 0.61).The status of stable MCI when compared with established AD was best predicted by a visual memory test, the "Doors" Test (F exit = 25). The second predictor was the RI48 Test (F exit = 15).
– "evolving" MCI vs. "stable" MCI (adjusted R^2 = 0.74). The status of evolving MCI when compared with stable MCI was best predicted by a visual memory test, the "Doors" Test (F exit=25). The second predictor was the RI48 Test (F exit = 14).

Published in: Journal of neurology (2005), vol. 252, iss. 1, pp. 47-55
Status: Postprint (Author's version)

Table 4 The baseline characteristics and performance on memory tests of stable versus evolving MCI patients

	Stable MCI	Evolving MCI[a]
Number	14	11
Age/years	71.0 (6.6)	71.8 (4.7)
CDR sum of boxes	1.6 (1.0)	2.2 (0.8)
MMSE initial/30	27.0 (1.2)	26.6 (2.2)
MMSE control/30	25.5 (2.1)	23.4 (2.9)
DRS/144	130.2 (6.4)	129.2 (5.4)
DRS memory/25	20.7 (2.8)	17.2 (3.9)
RI 48 Test		
• delayed recall/48	15.7 (3.5)	10.1 (4.2)
CERAD Word List Recall		
• sum of trials 1-3/30	15.6 (3.0)	13.7 (2.7)
• delayed recall/10	3.1 (2.8)	1.8 (1.3)
• savings (%)	44.5 (32.8)	33.5 (24.0)
"Shapes" Test		
• sum of trials 1-3/36	25.9 (7.5)	15.1 (8.3)*
• delayed recall/12	8.9 (2.7)	4.4 (3.1)*
• savings (%)	90.6 (13.9)	68.6 (22.8)
"Doors" Test		
• part A/12	9.9 (3.0)	6.6 (2.9)*
• part B/12	6.0 (3.1)	2.7 (1.5)

The values are mean(SD); [a] those patients initially diagnosed as MCI who fulfilled the criteria of probable AD at the follow-up; * statistically significant ($p < 0.05$) when the stable MCI were compared to evolving MCI

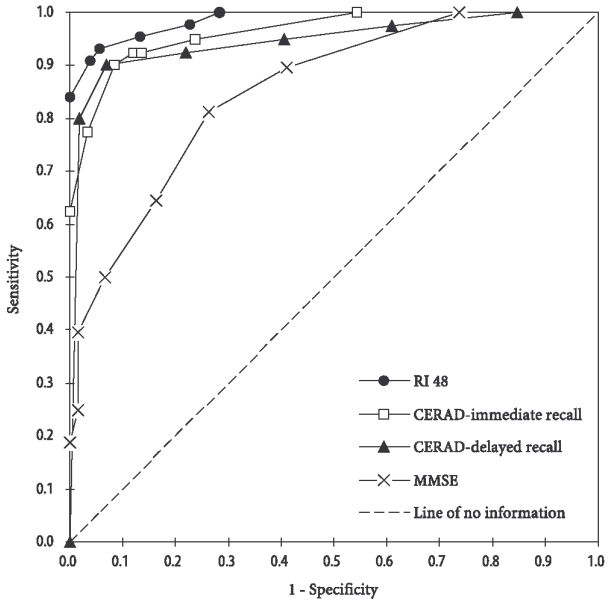

Fig. 1 ROC curves for the RI 48 Test, CERAD Test and the MMSE. The pathological group (MCI plus PrAD), labeled ptAD, was compared to the SMC. The AUC ± SE as well as the sensitivity%/specificity% for the best cut-off point, as resulted from the ROC analysis, are the following:
– RI 48 Test: 98.5 ± 0.9; 91/96 or alternatively 93/94
– CERAD immediate recall: 96.5 ± 1.6; 90/92
– CERAD delayed recall: 95.2 ± 2.4; 90/93
– MMSE: 85.3 ± 3.5; 81/74

Published in: Journal of neurology (2005), vol. 252, iss. 1, pp. 47-55
Status: Postprint (Author's version)

Discussion

The aim of this study was to determine the diagnostic value of the RI48 Test when compared with three established episodic memory tests in the early stage of AD. In order to target incipient AD cases, clinical criteria defining MCI [27] were used. Within the group of patients diagnosed as MCI at baseline, 44 % evolved to probable AD after 12–18 months. This conversion rate is higher than that previously reported for individuals classified using the same MCI criteria [27] adopted in this study. A possible reason for the higher conversion rate could be that some of the patients included in the MCI group might have had mild dementia.Although this is a possibility, it should be noted that the characteristics of our MCI group were comparable with those of the MCI group reported by Petersen [27]. The presence of only an isolated memory deficit in the MCI group was supported also by their normal performance in the comprehensive neuropsychological evaluation. It is more likely, however, that our finding might result from a bias of recruitment. Patients attending a Memory Clinic might be those with an accelerated evolution of their memory impairment, which makes themselves, their families or their general practitioners more keen to see k specialist advice. In community surveys both "fast"and "slow" deteriorating individuals would be detected, therefore lowering the overall evolution rate. Concerning the SMC group, it is worth noting that the relevance of subjective memory complaints as a risk factor for dementia is a controversial issue [9, 20], but the majority of authors acknowledge this symptom is not sufficient, if isolated, to predict future dementia. Our patients classed as SMC showed normal performance, virtually indistinguishable from that of age-matched controls, for memory as well as for all other cognitive functions. None of these individuals manifested signs of deterioration at follow-up.

With respect to the memory assessment, our hypothesis was that the cued recall test RI48 should allow a better separation of aging-related memory impairments from those impairments characteristic of AD, by boosting the performance of controls at a maximal level while giving little help to the AD patients, even when they are at the stage of MCI. The present study confirmed that the RI48 Test has a good sensitivity/specificity balance in the early diagnosis of AD. When comparing the RI48 Test with the other three well-established memory tests it is evident that there are significant differences between the tests for their diagnostic value in AD. The comparisons between immediate and delayed recall tests, between free and cued recall tests and between verbal and visual memory tests merit individual discussion.

It has been suggested that delayed recall of information [34] or alternatively the difference between immediate and delayed recall ("savings") [2] is typically the most impaired aspect of memory in AD patients. This study confirms that delayed recall measures (RI48 and the delayed recall from CERAD) are the most sensitive ones, although the immediate recall of CERAD was as sensitive as the delayed recall from the same test by the ROC method. In contrast, "savings" measures (from CERAD and the "Shapes" Test) are clearly less sensitive than both delayed recall and immediate recall measures. This finding is in agreement with the hypothesis that the memory deficit observed in AD is mainly one of encoding [12].

Free delayed recall is a difficult, effort requiring task, which implies not only a flawless encoding but also a strategic search in memory and the autogeneration of cues at retrieval, therefore deficits other than pure memory ones may account for poor performance at free recall. This is particularly true for elderly people, who are prone to atypical depression, sensory deficiencies, multimedication and may have concurrent debilitating diseases. The use of controlled cued memory tests would address this issue. It needs, however, to be clarified whether cued recall tests are as sensitive as free recall tests in detecting AD at the MCI stage. Recent evidence showed that free rather than cued recall of the FCSRT, a test based on a cueing technique similar to ours, is the best predictor of dementia over a 5-year follow-up [14]. However, the study by Grober [14] was based on a cohort survey and used different, less strict criteria for defining initially "non-demented" individuals. This study also employed different cognitive and functional scales in the evaluation of patients, therefore a direct comparison between these two studies is not possible. More importantly, the sample in that study was heterogeneous, including AD together with other types of dementia,which accounted for as much as 56% of the total number of patients who became demented. In our opinion, the conclusions of that study cannot be confidently applied to a typical AD population. Our study suggests, in contrast, that cued recall tests, such as the RI48 Test, are at least as good as free recall tests in the early diagnosis of AD, when having an adequate difficulty level.

The PrAD group showed impairment in all memory measures when compared with controls, whereas performance on some visual memory tests was spared in the MCI group. The initially labeled MCI patients who evolved to AD at follow up showed both low verbal and visual memory at initial assessment. In contrast, those MCI who remained "stable" at follow-up had impaired verbal memory but normal visual memory. This finding suggests a sequence in the appearance of memory deficits in AD, with a predominance of verbal memory deficits in the earliest stages.An alternative explanation might be that the verbal memory tests used in this study were more difficult than the visual ones. Using the same memory tests (with the exception of the RI48 Test), Greene et al. [12] found no clear-cut difference between verbal and visual memory impairment in AD. However, the

inclusion of a slightly more severe AD population in Greene's study could also account for the difference in results. An important finding of the present study that may be of clinical use is that when performance on visual memory tests similar to those used in this study is impaired, as well as on verbal memory tests, the risk of evolution to AD in the relatively short term is high.

CONCLUSIONS

This study showed that a memory test based on controlled encoding and retrieval conditions could be superior to tests based only on free recall in differentiating AD patients, including incipient cases, from healthy controls. Nevertheless, several limitations should be pointed out, such as the small and selected sample studied (exclusively patients attending a Memory Clinic) and the lack of "gold standard" other than the clinical diagnosis. Therefore, values of sensitivity and specificity reported in this study should be considered as relative, by comparison between the different tests used, rather than as absolute values for a particular test.

Acknowledgements: We are indebted to Anne Jacquemin, Genevieve Godfrind, Cecile Prairial, Mercedes Georges, Sophie Bechet and Delia Cicotti for helping us with data collecting. We thank Professor John Hodges for encouraging us all the way long and for his helpful advice in the conception of the paper. We thank Dr. Annalena Venneri for her useful comments on a draft version of this paper and for her help in improving the style. We are grateful to Janine Cooper for her amendments to the English. In addition, we thank the anonymous referees for their helpful remarks.

References

1. Adam S,Van der Linden M, Poitrenaud J, Kalafat M, et les membres du GRE-MEM (2004) L'épreuve de rappel in-dicé a 48 items (RI-48). In: Van der Linden M, Adam S, Agniel A, Baisset Mouly C, et al. (eds) L'évaluation des troubles de la mémoire: Présentation de quatres tests de mémoire épisodique (avec leur étalonnage). Marseille: Solal

2. Albert MS (1996) Cognitive and neu-robiological markers of early Alzheimer disease. Proc Natl Acad Sci USA 93:13547-13551

3. Baddeley A, Emslie H, Nimmo-Smith I (1994) Doors and People. Thames Valley Test Company

4. Buschke H, Sliwinski MJ, Kuslansky G, Lipton RB (1997) Diagnosis of early dementia by the Double Memory Test: Encoding specificity improves diagnostic sensitivity and specificity. Neurology 48:989-997

5. Buschke H (1984) Cued recall in amnesia. J Clin Exp Neuropsychol 6:433-440

6. de Partz MP, Bilocq V, De Wilde V, Seron X, Pillon A (2001) LEXIS: Tests pour 1'évaluation des troubles lexicaux chez la personne aphasique. Marseille: Solal

76. Elias MF, Beiser A, Wolf PA, Au R, White RF, D'Agostino RB (2000) The preclinical phase of Alzheimer disease. A 22-year prospective study of the Framingham Cohort. Arch Neurol 57:808-813

8. Ergis AM, Van der Linden M, Deweer B (1994) L'exploration des troubles de la memoire épisodique dans la maladie d'Alzheimer debutante au moyen d'une épreuve de rappel indicé. Rev Neuropsychol 4:47-68

9. Flicker C, Ferris SH, Reisberg B (1993) A longitudinal study of cognitive function in elderly persons with subjective memory complaints. J Am Geriatr Soc 41:1029-1032

10. Folstein MF, Folstein SE, McHugh PR (1975) "Mini Mental State". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189-198

11. Gifford DR, Cummings JL (1999) Evaluating dementia screening tests. Methodologic standards to rate their performance. Neurology 52:224-227

12. Greene JDW, Baddeley AD, Hodges JR (1996) Analysis of the episodic memory deficit in early Alzheimer's disease: evidence from the doors and people test. Neuropsychologia 34:537-551

13. Grober E, Buschke H, Crystal H, Bang S, Dresner R (1988) Screening for dementia by memory testing. Neurology 38:900-903

14. Grober E, Lipton RB, Hall C, Crystal H (2000) Memory impairment on free and cued selective reminding predicts dementia. Neurology 54:827-832

15. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29-36

16. Hogrefe CJ (1962) Test d2. Göttingen: Verlag für Psychologie

17. Hughes CP, Berg L, Danziger WL, Cohen LA, Martin RL (1982) A new clinical scale for the staging of dementia. Br J Psychiatry 140:566-572

18. Ivanoiu A, Adam S, Van der Linden M, et al. (2000) The episodic memory evaluation in incipient Alzheimer's disease. J Neurol 247(Suppl 3):III/13 (Abstract)

19. Ivanoiu A, Vanderlinden M, Seron X, Sindic CJM (1999) Contribution of free and cued recall to the early diagnosis of Alzheimer's disease. J Neurol 246 (Suppl 1):I/122 (Abstract)

20. Jonker C, Geerlings MI, Schmand B (2000) Are memory complaints predictive for dementia? A review of clinical and population-based studies. Int J Geriatr Psychiatry 15:983-991

21. Kertesz A, Mohs RC (1999) Cognition. In: Gauthier S (ed) Clinical Diagnosis and Management of Alzheimer's Disease, 2nd ed. London: Martin Dunitz, Ltd., pp 179-196

22. Mattis S (1988) Dementia Rating Scale: Professional manual. Odessa, FL: Psychological Assessment Resources

23. McKhann G, Drachman D, Folstein M, et al. (1984) Clinical Diagnosis of Alzheimer's disease. Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology 34: 939-944

24. Meiran N, Stuss D, Guzman A, Lafleche G, Willmer J (1996) Diagnosis of dementia. Methods for interpretation of scores of 5 neuropsychological tests. Arch Neurol 53:1043-1054

25. Morris JC, Mohs RC, Rogers H, Fillen-baum G, Heyman A (1988) Consortium to establish a registry for Alzheimer's disease (CERAD). Clinical and neuropsychological assessment of Alzheimer's disease. Psychopharmacol Bull 24:641-644

26. Petersen RC, Smith GE, Ivnik RJ, Kok-men E,Tangalos EG (1994) Memory function in very early Alzheimer's disease. Neurology 44:867-872

27. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment. Clinical characterization and outcome. Arch Neurol 56:303–308

28. Quality Standards Subcommittee of the American Academy of Neurology (1994) Practice parameter for diagnosis and evaluation of dementia (summary statement). Neurology 44: 2203–2206

29. Ritchie K, Fuhrer R (1992) A comparative study of the performance of screening tests for senile dementia using receiver operating characteristics analysis. J Clin Epidemiol 45:627–637

30. Rouleau I, Salmon DP, Butters N, Kennedy C, McGuire K (1992) Quantitative and Qualitative Analyses of Clock Drawings in Alzheimer's and Huntington Disease. Brain Cogn 18: 70–87

31. Small BJ, Fratiglioni L,Viitanen M, Winblad B, Backman L (2000) The course of cognitive impairment in pre-clinical Alzheimer disease. Three- and 6-year follow-up of a population-based sample. Arch Neurol 57:839–844

32. StatView (1992–1994) Abacus Concepts, Inc.

33. Tuokko H,Vernon-Wilkinson R,Weir J, Beattie BL (1991) Cued recall and early identification of dementia. J Clin Exp Neuropsychol 13:871–879

34. Welsh K, Butters N, Hughes J, Mohs R, Heyman A (1991) Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Arch Neurol 48:278–281