



This item was submitted to Loughborough's Institutional Repository by the author and is made available under the following Creative Commons Licence conditions.



**CC creative commons**  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

**BY:** **Attribution.** You must attribute the work in the manner specified by the author or licensor.

**Noncommercial.** You may not use this work for commercial purposes.

**No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>



## **The Hopkins Verbal Learning Test and screening for dementia**

Eva Hogervorst Marc Combrinck Pablo Lapuerta\* Judith Rue Kate Swales

Marc Budge

University of Oxford, Oxford Project to Investigate Memory and Ageing (OPTIMA), U.K.

\*Pharmaceutical Research Institute, Bristol-Myers Squibb, Princeton, NY

Robert Wood Medical School New Brunswick, NJ

Running head: The HVLT and screening for dementia (30 characters)

Dement Geriatr Cogn Disord; Dec 2002 13(1):13-20.

Corresponding author:

Dr. Eva Hogervorst

Oxford Project To Investigate Memory and Ageing

University of Oxford and Radcliffe Infirmary Trust,

Woodstock Road, Oxford, OX2 6HE, United Kingdom

Tel 00 44 1865 228512

Fax 00 44 1865 224099

Email: [eva.hogervorst@pharm.ox.ac.uk](mailto:eva.hogervorst@pharm.ox.ac.uk)

**Keywords:**

NINCDS/ADRDA - sensitivity-specificity - Hopkins Verbal Learning Test - Alzheimer's disease-dementia- verbal memory- MMSE – screening



**Abstract**

The present study investigated the sensitivity and specificity of the Hopkins Verbal Learning Test (HVLT) for demented patients (n=82, using NINCDS criteria) and 114 healthy controls - equivalent in age, years of education and gender-ratio - from the Oxford Project To Investigate Memory and Ageing. The HVLT 'Total recall' score had 87% sensitivity and 98% specificity for dementia using a cut-off score of 14.5. Using a 'Memory' score (the sum of the 'Total Recall' and the 'Discrimination Index') with a cut-off score of 24.5 gave a 91% sensitivity and 98% specificity for Alzheimer's disease cases when compared to controls. Unlike the MMSE, the HVLT has no ceiling effects and does not have to be adjusted for education. We conclude that the HVLT is an easy to administer, quick and well tolerated tool for the screening of dementia.

## **Introduction**

Dementia is characterised by a decline in memory and other cognitive functions, which impact on social and occupational performance. It is primarily a clinical diagnosis, which cannot be determined by laboratory findings alone [1]. When screening for dementia in large population studies and intervention trials, the tests used should ideally be short, reliable and easy to administer. The Hopkins Verbal Learning Test (HVLT) is claimed to fulfil these criteria. While most verbal memory tests are lengthy and complex, the HVLT is brief, is well tolerated by patients and was found to have no ceiling effects [2]. By comparison, the Cambridge Cognitive Examination (CAMCOG), a test battery developed to assess early dementia in older people [3] takes approximately 30 to 60 minutes to administer. An advantage of the HVLT over the widely used 5-10 minute Mini-Mental State Examination (MMSE) in the assessment of dementia [4] is its 6 equivalent or parallel forms, which allows serial testing with minimal learning effects [5-8]. In addition, the HVLT (like the CAMCOG) has a slightly higher sensitivity to detect mildly demented subjects than the MMSE [9].

The HVLT has been found to have a particularly high sensitivity and specificity for Alzheimer's disease (AD), the most common form of dementia [2, 10, 11]. The original validation study of the HVLT [2] investigated 45 subjects with probable AD, 3 amnesiacs and 18 healthy subjects aged 65 and older. They reported 94% sensitivity and 100% specificity for moderate to severe AD. A later study reported 96% sensitivity and 80% specificity for mild dementia (MMSE > 18) [9]. In studies of its discriminative capacity for different types of dementia, the HVLT was found to have a moderate (64%[10] to 77% [11]) correct classification rate for AD and Vascular Dementia (VaD).

However, most of these studies included very healthy elderly controls [2, 5] without any major co-morbidity, but did not apply the same criteria for the patient group. In addition, potential confounds like age, gender, education [2] and depression [2, 11], were not always adequately controlled for. These factors have been found to be risk

factors for both AD and poor memory function and therefore studies not controlling for these factors may have overestimated the discriminative capacity of the HVLТ. The present study further tested the specificity and sensitivity of the HVLТ, in terms of its ability to discriminate between controls and clinically diagnosed patients with dementia while controlling for these confounds. Diagnostic assessments were done by two independent experienced clinicians. In an earlier study, the agreement between our clinicians and the accuracy, specificity and sensitivity of clinical diagnostics when compared to histopathological criteria was found to be good [12].

## **Materials and methods**

### *Subjects*

This study was part of the ongoing Oxford Project To Investigate Memory and Ageing (OPTIMA) [13]. For the present study, 114 patients and 119 controls were included. Patients were usually referred to OPTIMA by their family physician from the Oxfordshire region because a dementia syndrome was suspected. Consecutive community dwelling, self-caring volunteer controls participated from the ongoing Foresight-Challenge study [14]. All subjects had undergone extensive medical screening at enrolment.

We excluded all cases without an informant, those with a MMSE score below 9 and those aged younger than 55 years at assessment (n=3). Controls were excluded when they scored lower than 24 on the MMSE or when they had a recent history of cerebrovascular disease (stroke, n=1). Co-morbidity or medication use was not a reason for exclusion, unless our clinicians considered it to interfere significantly with cognitive function (e.g. in case of alcohol abuses, see results section). Informed consent for all participants and ethical approval had been obtained prior to the study.



### *Diagnostic assessment*

Medical history, neurological examination, brain scans (CT or MRI and SPECT), and blood studies were reviewed for all participants. Diagnoses of normal functioning for age, Mild Cognitive Impairment (MCI) [15], Vascular Cognitive Impairment (VCI) according to the Hachinski criteria [16], possible or probable AD (NINCDS/ADRDA[1]), VaD (NINCDS/ AIREN [17]), other types of dementia, and psychiatric and neurological disorders (DSM-IV [18]) were made blinded to the HVLT performance of the subjects. Consensus meetings were held for ambiguous cases. In an earlier study, our clinicians had a substantial inter-rater reliability using NINCDS criteria (unweighted kappa=0.7). In addition, the clinical diagnosis of AD based on NINCDS/ADRDA criteria had 86% sensitivity and 79% accuracy when compared with the histopathological confirmed definite AD diagnosis (CERAD) [12].

Depression was taken into account by using the self-report depression scale of the Cambridge Examination for Mental Disorders of the Elderly (CAMDEX[3]). This examination also provided information about years of education. For patients, this information was obtained through an informant interview. For comparison, we included the MMSE score from the CAMDEX.

### *The Hopkins Verbal Learning Test (HVLT)*

The HVLT consists of a 12-item word list, composed of four words from each of 3 frequently printed semantic categories (e.g. 'precious stones'; 'human shelter'; 'animals') which subjects should recall in any order after the list has been read to them. This procedure is repeated 3 times. From this part of the test a total of the 3 free recall trials ('Total recall', cut-off score 10) is derived. Then, for a yes/no recognition, a list of 24 words is read, which consists of the 12 original words, 6 distractors from the same semantic categories (related false positives or FP-related) and 6 unrelated distractors (unrelated false positives or FP-unrelated). From this part of the test, the 'Discrimination index' (cut-off score 19) can be obtained.



### *Statistical analyses*

The HVLT provides two basic summary scores. The 'Total Recall' score is defined as the total frequency of correctly recalled items from all three free recall parts of the HVLT ( $=T1 + T2 + T3$ ). The recognition 'Discrimination Index' is calculated as (true positives – false positives). We also calculated 'Learning' which is the higher of either trial two or three minus the number recalled on trial 1. Lastly, a combination 'Memory' score was computed of the sum of 'Total Recall' and the 'Discrimination Index' to see if this would increase the discriminative ability of the test.

Demographic variables (age, gender-ratio, education in years, depression) were compared between groups using non-parametric Mann-Whitney and Chi square tests. We used Pearson's correlations to assess which characteristics were associated with dementia and the performance of the HVLT and MMSE subscores.

For the HVLT 'Total Recall', the 'Discrimination Index', the combination 'Memory' score and the MMSE, we produced Receiver Operating Characteristics (ROC) curves to examine the discriminative validity when screening for dementia. Cut-off scores were determined for maximum sensitivity and maximum (at least 98%) specificity. We also assessed whether the HVLT and MMSE variables had discriminative capacity for AD versus controls.

We performed logistic regression (backward conditional) using the cut off scores of the HVLT variables ('Total Recall' and the combination 'Memory' score) with the highest specificity (as obtained with the ROC curve analyses) as the dependent variables. In these stepwise analyses, we included dementia (yes/no) and potential confounds such as age, gender, education and depression as independent variables. These analyses tested whether cut-off scores for the HVLT needed to be modified according to patient characteristics (age, education, gender, depression). For comparison, we did the same for the MMSE. All analyses were conducted in SPSS.

## Results

### *Subjects*

Of the 119 controls screened, 5 were excluded (1 with VCI, 3 were psychiatric cases and 1 had a neurological disorder). Of the 111 patients originally referred as demented, 29 were excluded (9 had MCI, 5 VCI and 4 were controls, 9 were psychiatric cases and 2 subjects had a neurological disorder).

We thus included 114 controls (which included 8 subjects with MCI) and 82 mild to moderate demented cases (MMSE 9-28) in the analyses. Of the demented group, 68 (82%) were diagnosed as AD (29 possible and 31 probable AD and 8 mixed AD/VaD cases), 6 as VaD, 3 as Lewy Body Dementia (LBD) and 5 as other types of dementia (ODS, such as frontotemporal dementia).

### *Demographics*

The demographic characteristics of the demented (AD, VaD, mixed, LBD, ODS) versus control participants (controls, MCI) were compared using Mann Whitney tests and Chi square (for gender ratio). There was no difference in age ( $Z=-1.17$ ,  $p=0.24$ ; mean age 75 years SD 7), gender ratio (49% females in controls, 42% in demented  $\chi^2=1.29$ ,  $p=0.23$ , and 46% in AD,  $\chi^2=0.30$ ,  $p=0.58$ ), or years of education ( $Z=-1.47$ ,  $p=0.14$ , mean 11 years SD 2) between groups. There was a significant difference between groups in CAMDEX derived depression scale scores ( $Z=-5.49$ ,  $p<0.000001$ ,  $2 \pm 2$  in controls Vs  $4 \pm 3$  in the demented groups on average) but all scores were still within the normal range (0-11). With regard to demographic characteristics, there was no difference (compared to the demented group overall) when the AD group was compared to controls separately.

Pearson's correlations showed the dementia group had a lower performance than controls on the 'Total Recall' ( $r=-0.81$ ,  $p < 0.0001$ ), the 'Learning' score ( $r=-.43$ ,  $p < 0.0001$ ), the 'Discrimination Index' ( $r=-.66$ ,  $p < 0.0001$ ), the combination 'Memory' score ( $r=-.82$ ,  $p < 0.001$ ) and the MMSE ( $r=-0.80$ ,  $p < 0.0001$ ) (see also table 1). The MMSE performance was associated with the 'Total Recall' ( $r=0.79$ ,  $p < 0.0001$ ), the 'Learning' score ( $r=0.44$ ,  $p < 0.0001$ ), the 'Discrimination Index' ( $r=0.74$ ,  $p < 0.0001$ ), the combination 'Memory' score ( $r=0.84$ ,  $p < 0.0001$ ).

Older subjects had a slightly lower 'Total Recall' ( $r=-.16$ ,  $p < 0.05$ ) and combination 'Memory' score ( $r=-.15$ ,  $p < 0.05$ ). A scatter plot revealed a slight decrease in 'Total Recall' performance with age but this was mainly apparent for controls. The percentage of variance explained by age on this measure for controls was larger ( $R^2=0.03$ ) than for cases ( $R^2=0.007$ ). For example, at 60 years of age, controls had an average performance on the 'Total Recall' of 24.5, while cases scored 11.5 at that age. At 90 years of age, controls had an average 'Total Recall' performance of 21 (difference=3.5), while cases scored 10 (difference=1.5).

A higher MMSE (but not 'Total Recall') was associated with having had more years of education ( $r=0.16$ ,  $p=0.02$ ). Scatter plots (fig. 1) showed that this effect was mainly apparent for the demented cases on the MMSE. The percentage of variance of the MMSE performance explained by education was higher for the cases ( $R^2=0.04$ ) than for controls ( $R^2=0.004$ , see fig. 1).

Members of the dementia group were more likely to have a higher depression score ( $r=.40$ ,  $p < 0.0001$ ). The depression score also increased with age ( $r=.15$ ,  $p < 0.05$ ), and was associated with a lower performance on all the tests, the MMSE ( $r=-.21$ ,  $p < 0.0001$ ), the 'Total Recall' ( $r=-.31$ ,  $p < 0.0001$ ), 'Learning' ( $r=-.19$ ,  $p < 0.001$ ), the 'Discrimination Index' ( $r=-.21$ ,  $p < 0.005$ ) and the 'Memory' score ( $r=-.31$ ,  $p < 0.0001$ ). There were no associations with gender and the cognitive tests.

- insert fig 1-

The ROC curves were produced by plotting the sensitivity against the 1-specificity for each score on the HVLT 'Total Recall' in discriminating between demented cases versus controls (fig. 2) and between AD cases versus controls. In table 2, a summary of the ROC curve analyses for detecting dementia and AD is shown with the cut-off scores (and corresponding sensitivity and specificity) for the HVLT 'Total Recall', 'Discrimination Index', combination 'Memory' score and the MMSE. The 'Total Recall' showed a slight trend for a better sensitivity than the MMSE using a cut-off score of 14.5. The 'Total Recall' and the combination 'Memory' score had a slightly lower specificity with similar sensitivity when compared to the MMSE but this depended on the cut-off score used (e.g. 100% sensitivity favoured the HVLT parameters). The combination 'Memory' score gave a better overall discrimination for AD patients than the MMSE. Using a cut-off score of 24.5 and 98% specificity for both tests, rendered 91% sensitivity for the combination 'Memory' score and 85% for the MMSE (table 2).

-insert table 2-

-insert fig 2-

Stepwise backward conditional logistic regression entering age, gender, education, depression and dementia (y/n) showed that for the HVLT 'Total Recall' score (recoded as below '0' and equal to or above '1' the cut-off score of 14.5) dementia was the only significant predictor of performance (correct classification of 94%). Similar results were found for the combination 'Memory' score (recoded as below '0' and equal to or above '1' the cut-off score of 24.5, correct classification 93%). For the MMSE (using a cut-off score of 24.5, correct classification of 91%) belonging to the dementia category was a significant predictor of performance. There were also trends for both education ( $p=0.10$ ) and depression ( $p=0.07$ ) in predicting performance below the cut-off of 24.5 of the MMSE. Using AD instead of dementia as an independent variable, gave for the 'Total Recall' and the 'Memory' score a 95% and for the MMSE a 93% correct classification score. No other variables were entered.

## **Discussion**

We investigated the sensitivity and specificity of the HVLT, in terms of its ability to distinguish between controls and cases with mild to moderate dementia and its ability to discriminate between patients with Alzheimer's disease (AD) and controls. We found that the 'Total Recall' with a cut-off score of 14.5 discriminated well between patients and controls. The 'Discrimination Index' by itself did not add much in terms of its discriminative capacity but in combination with the 'Total Recall' (the 'Memory' score), it was seen to give a good classification rate, especially for AD. The comparison of the HVLT parameters with the MMSE, in terms of its sensitivity and specificity is limited by our selection criteria for controls (MMSE had to be higher than 24).

However, the distributions of the HVLT 'Total Recall' and MMSE were clearly different. A ceiling effect was evident for the MMSE, with 30% of controls having perfect scores. No controls scored perfectly on the HVLT (fig.3). In addition, the MMSE performance was better with higher education, which may require adjustment of the individual scores. These factors make the HVLT more useful for screening purposes.

Depression scores (within the normal range) were associated with a poorer performance on all the cognitive test parameters. Interestingly, most of the psychiatric cases (n=9), who were excluded from analyses, did not perform below the cut-off point of 14.5 on the 'Total Recall'. Three cases performed below the cut-off score, all of whom had complex histories of treatment resistant depression (with ECT), substance abuse and suspected organic brain syndrome. In addition, although depression scores were higher (but again within the normal range) in the demented cases, depression did not explain the dementia-related variance of HVLT test performance. It should also be kept in mind that these data, unlike those for the controls, were obtained from informants, which may not always be accurate (i.e. informants may overestimate depression, whilst controls could be less likely to report a depressed mood).

Our findings are largely in concordance with the results of earlier studies which reported a slightly higher discriminative capacity of the HVLT than ours. A closer look at the other studies [2, 5, 9, 11] revealed that our controls had a slightly lower performance than theirs. This may be explained by their very rigorous health screening of controls [2, 5] or the lack of control for age, education, gender [2] and depression [2, 11]. These factors may have overestimated the normal performance of elderly subjects in the previous studies. In our study, controls were only excluded if disease was thought to be implicated in cognitive dysfunction, not a priori. In addition, we included subjects with MCI which gives a more realistic reflection of the elderly clinical population who are not demented. Our cut-off score of 14.5 for 'Total Recall' was higher than the original 10 as suggested by Brandt [2]. Our AD patients were in the early stages and had high MMSEs, which may have made the difference between groups smaller. Our 'Discrimination Index' also showed a lower discriminative capacity compared to the other studies. Craik and Byrd [19] have shown that elderly subjects perform poorly on tasks with little environmental or contextual support. Recognition has more contextual support than recall, which makes it less sensitive to earlier stages of dementia.

There was a difference in the discriminative capacity of the HVLT variables for controls and dementia overall and AD by itself. Box plots (fig. 4) showed that the VaD patients (who had been included in the dementia group) had a higher performance than any of the other types of dementia groups, including AD, particularly on the 'Total Recall'. Overall, VaD patients had a lower performance than the controls but performed similarly to the MCI, VCI, psychiatry and neurological cases. This may reflect the general difficulty in distinguishing the latter categories from dementia [20, 21]. However, numbers in these groups were considered too small to draw conclusions even though post hoc analyses showed that the differences were statistically significant. It is important to be able to discriminate between different types of dementia

early in the disease as they may require different treatment strategies. Other studies have found that verbal memory tests with list recall, like the Californian Verbal Learning test, can discriminate between VD and AD [22]. However, one of the major criticisms is that these tests usually take a long time to administer. In contrast, the HVLT is easy to administer, quick and is well tolerated by participants. Most importantly, the HVLT does not have ceiling effects and is not sensitive to educational levels (unlike the MMSE) and hence needs no adjustment for the individual scores.

-insert fig 4-

## **Conclusion**

We have found that the HVLT discriminated well between demented cases and controls and is a useful tool in clinical and epidemiological practice. The cut-off score of 14.5 of the HVLT 'Total recall' score was shown to give a good discrimination between cases and controls. In this study we were mainly concerned with obtaining the highest specificity (> 98%, i.e. correctly identifying controls). If the sensitivity needs to be higher for research purposes, a higher cut-off for the 'Total Recall' of 19.5 or the use of the 'Memory' score with a cut-off point of 24.5 is advisable.

## **Acknowledgements**

The authors would like to thank the members - in particular Prof. A. David Smith for his valuable comments - and participants of OPTIMA for making this study possible. This work was supported by a grant from Bristol-Myers Squibb.

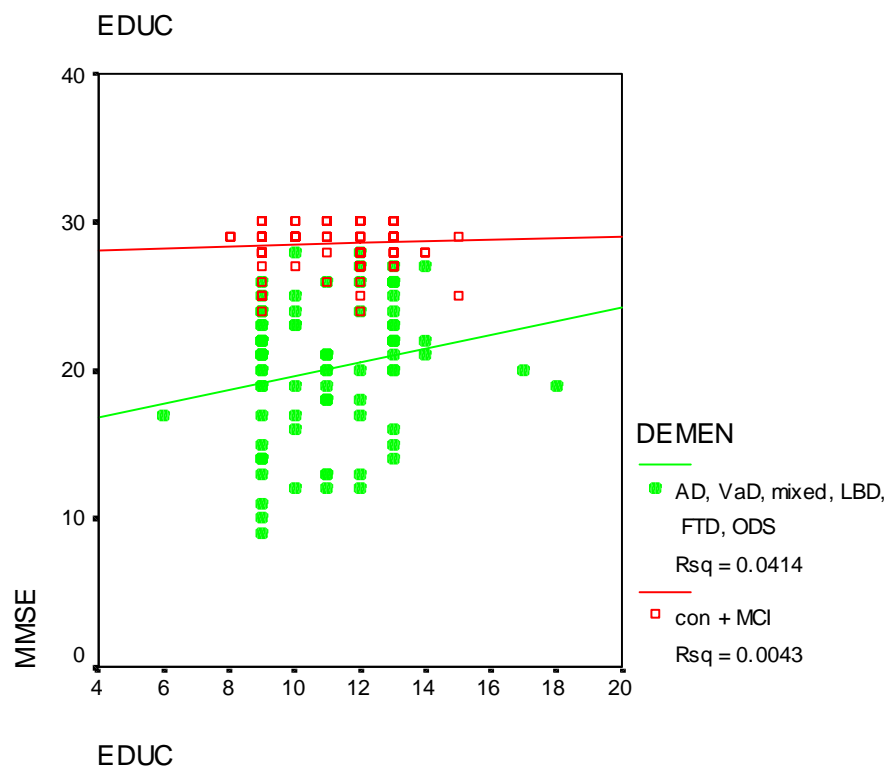
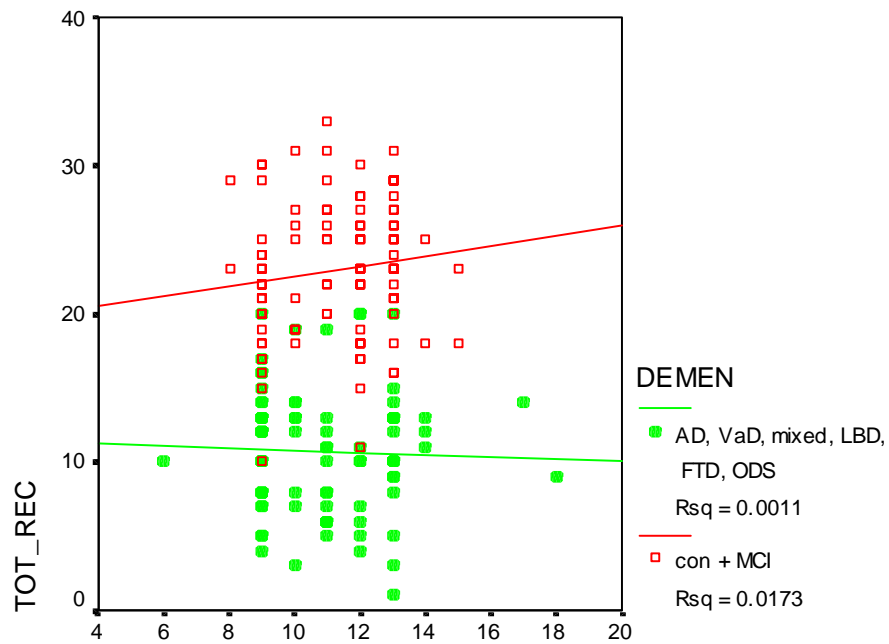
## References

- 1 McKhann G, Drachmann 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 Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34: 939-944.
- 2 Brandt J: The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clin Neuropsychol* 1991; 5: 125-142.
- 3 CAMDEX: The Cambridge Examination for Mental Disorders of the Elderly. Cambridge, Cambridge University Press, 1988.
- 4 Folstein M, Folstein S, McHugh P: 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatry Res* 1977; 12: 189-198.
- 5 Benedict RHB, Schretlen D, Groninger L, Brandt J: Hopkins Verbal Learning Test--Revised: Normative data and analysis of inter-form and test-retest reliability. *Clin Neuropsychol* 1998; 12: 43-55.
- 6 Krebs R: The Hopkins Verbal Learning Test: An alternative to the MMSE? *Gerontologist* 1994; Oct;34: 692.
- 7 Rasmusson DX, Bylsma FW, Brandt J: Stability of performance on the Hopkins Verbal Learning Test. *Arch Clinic Neuropsychol* 1995; Jan-Feb;10: 21-26.
- 8 Carpenter BD, Strauss ME, Ball AM: Telephone assessment of memory in the elderly. *J Clinic Geropsychol* 1995; Apr; 1: 107-117.
- 9 Frank RM, Byrne GJ: The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia. *Int J Geriatr Psychiatry* 2000; 15: 317-324.
- 10 Shapiro AM, Benedict RHB, Schretlen D, Brandt J: Construct and concurrent validity of the Hopkins Verbal Learning Test-Revised. *Clin Neuropsychol* 1999; 13: 348-358.
- 11 Barr A, Benedict R, Tune L, Brandt J: Neuropsychological differentiation of Alzheimer's disease from vascular dementia. *Int J Geriatr Psychiatry* 1992; 7: 621-627.
- 12 Hogervorst E, Barnetson L, Jobst KA, Nagy Z, Combrinck M, Smith AD: Diagnosing dementia: interrater-reliability assessment and accuracy of the NINCDS?ADRDA criteria versus CERAD histopathological criteria for Alzheimer's disease. *Dement Geriatr Cogn Disord* 2000; 11: 107-113.
- 13 Clarke R, Smith, A.D., Jobst, K.A., Refsum, H., Sutton, L., Ueland, P.M.: Folate, vitamin B12 and serum

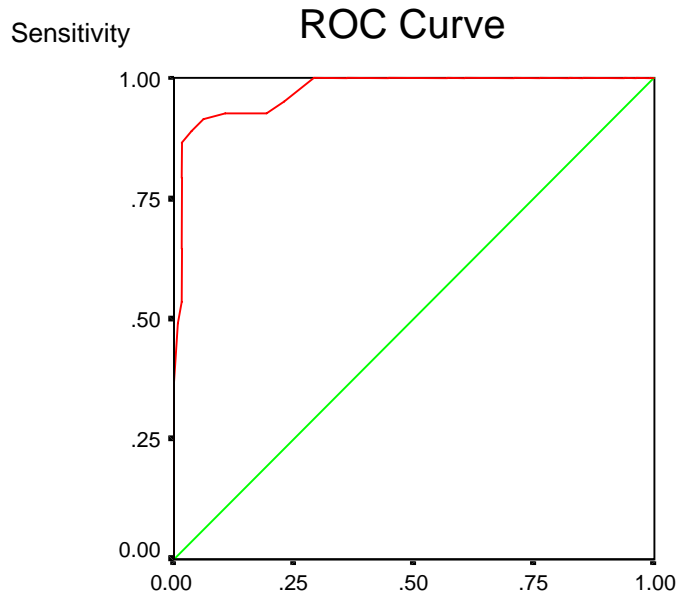


- homocysteine as candidate risk factors for confirmed Alzheimer's disease. *Arch Neurol* 1998; 55: 1449-1455.
- 14 Budge M, Johnston C, Hogervorst E, DeJager C, Milwain E, Iversen SD, Barnetson L, King E, Smith AD: Plasma total homocysteine and cognitive performance in a volunteer elderly population. *Ann NY Acad Sci* 2000; 903: 407-411.
- 15 Petersen RC, Smith GE, Waring SC, Ivink RJ, Tangalos EC, Kokmen E: Mild cognitive impairment. *Arch Neurol* 1999; 56: 303-308.
- 16 Wentzel C, Darvesh S, McKnight C, Shea C, & , Rockwood K: Inter-rater reliability of the diagnosis of Vascular Cognitive Impairment at a memory clinic. *Neuroepidemiology* 2000; 19: 186-193.
- 17 Gold G, Giannakopoulos P, Montes-Paixao C, Herrmann FR, Mulligan R, Michel JP, Bouras C: NINCDS/AIREN criteria. *Neurology* 1997; 49: 692.
- 18 Diagnostic and Statistical manual of Mental disorders, DSM-IV. Washington D.C., American Psychiatric Association, 1994.
- 19 Craik FIM, Bird, M.: Ageing and cognitive deficits the role of attentional resources.; in F. I. M. Craik S. Trehub (eds): *Aging and Cognitive Processes*. New York, Plenum, 1982, pp 191-211.
- 20 Verhey FRJ, Jolles J, Ponds RWHM, Rozendaal N, Plugge LA, Vreeling FW, van der Lugt PJM: Diagnosing dementia: a comparison between a monodisciplinary and a multidisciplinary approach. *J Neuropsychiat Clin Neurosc* 1993; 5: 78-85.
- 21 Coen RF, O'Mahoney D, Bruce I, Lawlor BA, Walsh JB, Coakley D: Differential diagnosis of dementia: a prospective evaluation of the DAT Inventory. *J Am Geriatr Soc* 1994; 42: 16-20.
- 22 Looi JCL, Sachdev PS: Differentiation of vascular dementia from AD on neuropsychological tests. *Neurology* 1999; 53: 670-678.

Fig.1. Scatter plots of HVLT 'Total Recall' (TOT\_REC) and MMSE versus education (EDUC)



**Fig. 2.** ROC curve of the sensitivity and specificity of the HVLT' Total Recall' in demented vs controls



1 - Specificity

Diagonal segments are produced by ties.

**Fig. 3.** Scatter plot of the HVLT 'Total Recall' against the MMSE with different cut-off scores

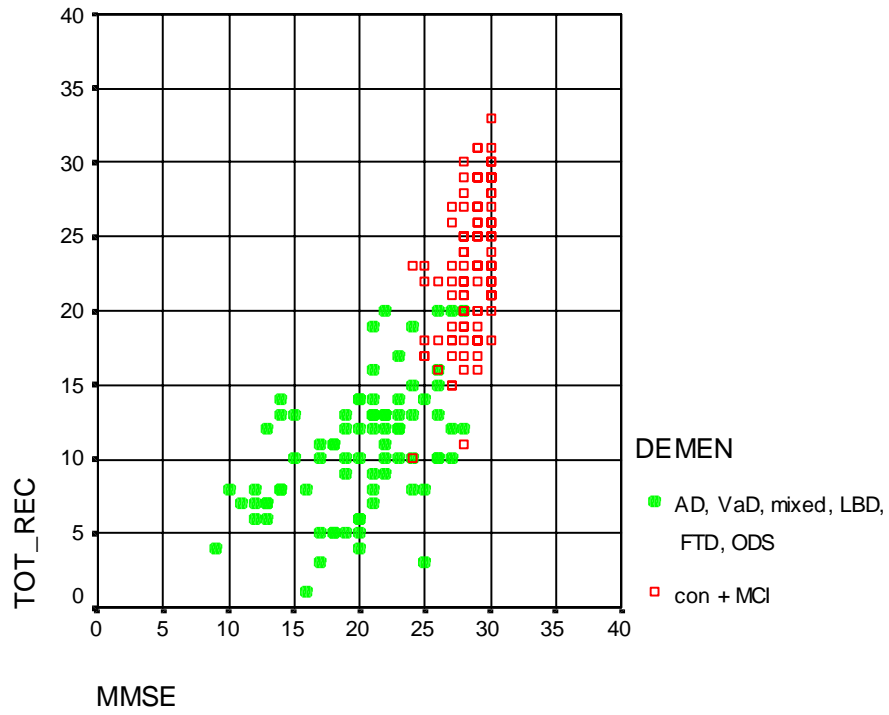
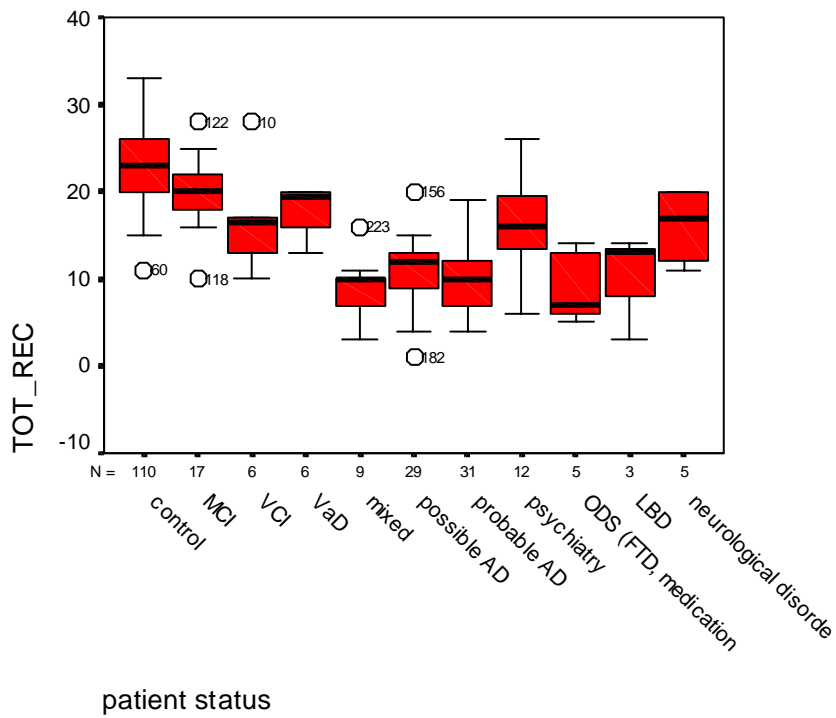


Fig. 4.

Box plot of the median performance of different diagnostic groups on the HVLT 'Total Recall'



**Table 1.** The performance of the different groups on the individual HVLT trials (T1 to T3), the 'Total Recall'; 'Learning' the 'Discrimination index', the 'Memory' score and the MMSE

Variable	Controls	Demented overall	AD (separate)
HVLT			
T1	5.95 (1.57)	2.73 (1.46)***	2.59 (1.39) ***
T2	7.84 (1.85)	3.68 (1.57) ***	3.52 (1.42) ***
T3	9.20 (1.90)	4.30 (1.81) ***	4.13 (1.59) ***
Learning	3.27 (1.72)	1.76 (1.40) ***	1.74 (1.39) ***
Total Recall	22.99 (4.49)	10.72 (4.24) ***	10.24 (3.70) ***
Discrimination Index	10.84 (1.07)	6.99 (3.08) ***	6.65 (3.07) ***
False Positives-related	.84 (.93)	2.12 (1.46) ***	2.24 (1.46) ***
False Positives-unrelated	.53 (.26)	1.20 (1.44) ***	1.35 (1.48) ***
Combination 'Memory' score	33.83 (4.88)	17.71 (6.24) ***	16.88 (5.62) ***
MMSE	28.52 (1.45)	20.04 (4.60) ***	25.27 (5.16) ***
N	114	82	68

\*\*\*= Mann-Whitney U test asymp. significance  $p < 0.00001$  (2-tailed)

**Table 2.** Specificity and sensitivity of different cut-off points for the HVLT and MMSE in discriminating between controls and demented and AD as determined by ROC curves

Area	Std. Error	95% C.I.	Sign.	Cut-off	Sensitivity	Specificity
<i>Total recall: demented – controls</i>						
0.97	0.01	0.95 - 0.99	0.0001	14.5	87%	<b>98%</b>
				19.5	95%	77%
				20.5	<b>100%</b>	71%
<i>Total recall: AD – controls</i>						
0.98	0.008	0.97 - 0.99	0.0001	14.5	91%	<b>98%</b>
				19.5	<b>99%</b>	77%
				20.5	<b>100%</b>	71%
<i>Discrimination Index: demented – controls</i>						
0.88	0.03	0.83 - 0.94	0.0001	7.5	51%	<b>100%</b>
				11.5	94%	34%
<i>Discrimination Index: AD – controls</i>						
0.9	0.03	0.85 - 0.96	0.0001	7.5	57%	<b>100%</b>
				11.5	94%	34%
<i>Memory: demented – controls</i>						
0.98	0.01	0.96 - 0.99	0.0001	24.5	84%	<b>98%</b>
				30.5	<b>99%</b>	75%
				31.5	<b>100%</b>	69%
<i>Memory: AD – controls</i>						
0.99	0.01	0.98 - 0.99	0.0001	24.5	91%	<b>98%</b>
				28.5	<b>99%</b>	86%
				30.5	<b>100%</b>	75%
<i>MMSE: demented – controls</i>						
0.98	0.01	0.97 - 0.99	0.0001	24.5	83%	<b>98%</b>
				27.5	<b>98%</b>	82%
				28.5	<b>100%</b>	57%
<i>MMSE: AD – controls</i>						
0.98	0.01	0.97 - 0.99	0.0001	24.5	85%	<b>98%</b>
				27.5	<b>99%</b>	82%
				28.5	<b>100%</b>	57%

Headings of figures and tables

**Fig.1.** Scatter plots of HVLT 'Total Recall' (TOT\_REC) and MMSE versus education (EDUC)

**Fig. 2.** ROC curve of the sensitivity and specificity of the HVLT' Total Recall' in demented vs controls

**Fig. 3.** Scatter plot of the HVLT 'Total Recall' against the MMSE with different cut-off scores

**Fig. 4.**

Box plot of the median performance of different diagnostic groups on the HVLT 'Total Recall'

**Table 1.** The performance of the different groups on the individual HVLT trials (T1 to T3), the 'Total Recall'; 'Learning' the 'Discrimination index', the 'Memory' score and the MMSE

**Table 2.** Specificity and sensitivity of different cut-off points for the HVLT and MMSE in discriminating between controls and demented and AD as determined by ROC curves