Demenzdiagnostik mit Hilfe der Vienna Neuropsychologischen Testbatterie (VNTB): Standardisierung, Normierung und Validierung

The Vienna Neuropsychological Test Battery (VNTB) for detecting Alzheimer's Dementia: standardization, norms, and validation

Johann Lehrner, Joachim Maly, Andreas Gleiß, Eduard Auff & Peter Dal-Bianco

Zusammenfassung

Die frühzeitige Erkennung der Alzheimerkrankheit wird aufgrund von neuen pharmakologischen Therapieoptionen immer wichtiger. Ziel der vorliegenden Studie war die Standardisierung und Normierung der Vienna Neuropsychologischen Testbatterie (VNTB). Zusätzlich sollte die diagnostische Wertigkeit der eingesetzten neuropsychologischen Testverfahren für die Diagnose der Alzheimerkrankheit überprüft werden. In die Studie eingeschlossen wurden Patienten, die in der Gedächtnisambulanz der Neurologischen Universitätsklinik mit Gedächtnisstörungen vorstellig wurden. 136 Patienten wurden klinisch untersucht und unterzogen sich einer ausführlichen neuropsychologischen Untersuchung. 78 Patienten erhielten die Diagnose Alzheimer Demenz und 58 Patienten wurden als kognitiv nicht beeinträchtig eingestuft. Sensitivität, Spezifität, positiv prädiktiver Wert und negativ prädiktiver Wert wurde für alle Variablen erhoben. Die neuropsychologischen Variablen der VNTB konnten statistisch signifikant zwischen der kognitiv nicht beeinträchtigen Gruppe und der Alzheimergruppe differenzieren. Die vorliegende Studie liefert Hinweise für die Brauchbarkeit der VNTB für die Diagnose der Alzheimerkrankheit.

Abstract

Early detection of dementia is becoming more and more important due to the advent of pharmacologic treatment. The goals of this study were to report standardization procedures and norms for the Vienna Neuropsychologische Testbatterie (VNTB) and further to evaluate the diagnostic utility of the used psychometric measures. Patients complaining about memory problems and who came to the memory outpatient clinic for assessment of their memory disorder were included in the study. One hundred thirty-six patients underwent a clinical examination and completed a battery of standard cognitive tests at study entry. Seventy-eight received the diagnosis of Alzheimer's disease and fifty-eight were categorized as cognitively unimpaired controls. Sensitivity, specificity, receiver operating characteristics (area under curve, AUC), positive predictive value and negative predictive value for each neuropsychological test were determined. All neuropsychological variables significantly separated dementia patients and controls on a group basis. Receiver operating characteristics based on the measure of AUC of the specific neuropsychological tests ranged from 0.78 to 0.99. Our study contributes knowledge regarding the diagnostic value of the VNTB in patients with memory impairment attending a memory clinic.

1. Introduction

In recent years the advent of pharmacological treatment for Alzheimer's Disease (AD) has spurred the interest of diagnosing dementia as early as possible in order to provide early treatment. Published criteria for AD diagnosis such as those developed by the Work Group of the National Institute of Neurological and Communication Disorders and the Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) (McKhann et al., 1984) require standard assessment of patients.

The neuropsychological diagnosis is based on impairments in relevant cognitive domains such as psycho-motor skills, attention, language, memory and executive functions. These domains can be assessed by commonly used cognitive measures. It is likely that more than one measure may be required to tap complex constructs such as psycho motor skills (eye-hand coordination), attention (selective attention, divided attention), language (confrontation naming, verbal fluency), memory (learning and immediate and delayed recall of verbal material) and executive function (planning, concept formation, shifting cognitive sets).

One problem in the field of early dementia detection is the reliable discrimination of normal aging from mild dementia particularly in high functioning individuals and people with life-long poor cognitive functioning. Several screening measures such as the MMSE (Folstein et al., 1975), clock drawing test (Powlishta et al., 2002) and seven minute screen (Meulen et al., 2004) as well as neuropsychological test batteries such as the CERAD neuropsychological test battery (Chandler et al., 2005) and the neuropsychological test battery for the Alzheimer's Disease Cooperative Study (Grundman et al., 2004) are available.

Because no commonly accepted specific test battery is recommended, a recent review of early detection of dementia with recommendations for future research recently published advocated additional research to develop appro-

priate and sensitive neuropsychological methods (Luis et al., 2003).

Thus, the goals of the reported study were two-fold: First, we wanted to report standardization procedures and norms for the newly established neuropsychological test battery for detecting dementia, and second, we aimed to investigate convergent validity and diagnostic value of neuropsychological testing in diagnosing early dementia. Sensitivity, specificity, receiver operating characteristics (area under curve; AUC), positive predictive value and negative predictive value for each neuropsychological test were determined.

2. Methods

2.1. Subjects

Patients complaining of memory problems who came to the memory outpatient clinic for assessment of their memory disorder were included in the study. The study protocol was in accordance with the Helsinki Declaration. Patients were either referred by physicians or were self-referrals. The included area of the study was Vienna. All patients received a complete neurological examination, standard laboratory blood tests and psychometric testing. In most cases a CT scan or MRI scan of the brain was obtained. Electroencephalogram and single-photon emission computed tomography scans were performed on some patients. In determining significant cerebrovascular disease, both neuroimaging and clinical patient features were used.

Inclusion and exclusion criteria were similar to other studies. Patients were excluded from the study if any of the following conditions applied (a) evidence of cortical stroke as determined by neuroradiologic examination, (b) history of severe head injury, (c) current psychiatric diagnosis according to ICD-10, (d) any medical condition that lead to severe cognitive deterioration including renal, respiratory, cardiac and hepatic disease, (e) less than 50 years of age, (f) non-AD dementias such as Frontotemporal dementia or Lewy body disease. After the completion of the evaluation, a consensus committee meeting was held involving the neurologist, neuropsychologists and other study personnel who had evaluated the patients. Diagnosis of dementia was made according to the DSM-IV and the Alzheimer's Disease (AD) subjects were diagnosed using the NINCDS-ADRDA guidelines. All AD patients had MMSE score equal or less 24 (Folstein et al., 1975).

One hundred thirty-six patients fulfilled the criteria and were thus included in the study. Seventy-eight patients received the diagnosis of AD dementia and fifty-eight patients were categorized as controls without significant cognitive impairment. Mean age of AD patients was 74.6 ± 6.9 years and mean age of controls was 67.2 ± 8.2 years. There were 43.6% males and 56.4% females in the dementia group and 39.7% males and 60.3% females in the control group. Mean years of formal schooling was 9.6 ± 2.7 for dementia patients and mean years of formal schooling was 11.7 ± 3.7 for the controls. Mean MMSE performance of patients was 21.1 ± 2.0 and mean MMSE performance of controls was 28.0 ± 1.4.

2.2. Neuropsychological Measures

All patients were subjected to a battery of neuropsychological tests that included the domains psycho motor speed, concentration/attention, language, memory and executive functioning (Lezak, 1995). Cognitive function tests were selected to assess a broad range of cognitive abilities commonly affected by AD and other dementias. Psycho motor speed was assessed using the symbol counting task from the cerebral insufficiency test (C.I.) (Lehrl & Fischer, 1997) and the Trail Making Test A (Reitan, 1979).The Alters-Konzentrations-Test (AKT) (Gatterer, 1990), a geriatric cancellation test, the digit symbol subtest of the German WAIS-R (Tewes, 1994) and the Trail Making Test B (Reitan, 1979) were applied to assess attention. In order to test language functions, we used verbal fluency tasks and a confrontation naming task (Goodglass, 1983). Naming as many animals, supermarket items and tools that came to mind within one minute for each task was used tap semantic verbal fluency. Naming as many words beginning with the letters b, f, l that came to mind within one minute for each task was used to tap phonematic verbal fluency. The modified Boston Naming Test (BNT) (Morris et al., 1989) was used for assessing naming capabilities. Episodic memory was tested using the Verbal Selective Reminding Test (VSRT) (Lehrner et al., 2006) which is the Austrian paper pencil version of the Memory Assessment Clinics (MAC) (Steiner, 1998) Grocery List Selective Reminding Test with the subtests of immediate recall, total recall and delayed recall (Crook et al., 1986; Youngjohn et al., 1991). Executive functions were investigated using the difference score of the Trail Making Test A and B (Reitan, 1979), the Five-Point Test (Regard et al., 1982), the Maze Test and the Stroop Test from the Nuremberg Aging Inventory (NAI) (Oswald & Fleischmann, 1997) and the interference test from the C.I. (Lehrl & Fischer, 1997). Cognitive testing for each patient lasted approximyately 60 minutes.

Adequate normative data for the above mentioned measures were available. Cognitively normal subjects were drawn from the General Hospital of Vienna including the Department of Neurology (Memory Clinic). The cognitively normal subjects underwent a standard medical evaluation and were judged as being in good health. Criteria for normal function were identified similar to those in the Mayo research studies (Ivnik et al., 1992; Petersen et al., 2004): (a) no active neurological or psychiatric disease, (b) no psychotropic medications, and (c) the subjects may have medical disorders but neither they nor their treatment compromises cognitive function. Enrolled subjects of the cognitively normal sample were required to have an MMSE score greater than or equal to 27 and a memory score greater than -1.5 standard deviations on the faces recognition test of the MAC test battery (Crook et al., 1986).

2.3. Statistical methods

Because age, education and gender effects on cognitive variables have been reported in the literature (Chandler et al., 2005), regression based norms (z-scores) were calculated using a multiple linear regression formula for each single subtest. The regression coefficients and other parameters were calculated in the cognitively normal sample as described below. Demographic variables of age, gender and education were used as regressors (adjustment variables). Thus, z-scores for a specific subtest are reported as the standardized residuals of an appropriate regression model (Berres et al., 2000) (for details see below) and indicate the relative degree of impairment from normal in SD units, thereby allowing comparison across different cognitive tests. Due to administrative reasons, the number of normative subjects on which the corresponding regression is based differs and ranges from 122 to 434 (see table 1).

First, appropriate transformations for the subtest variables were selected (logarithmic, square root, arcsine or power transformations) in order to provide sufficient approximation of the regression residuals to a normal distribution (after adding, if necessary, a suitable constant to obtain non-negative values).

The variables considered for adjustment in the regression models were age, years of education and gender, as well as quadratic and cubic terms and interaction terms of each pair of demographic variables. For the calculation of the higher order and the interaction terms the respective variables were centred at their sex specific median. The variables for adjustment were chosen on the basis of Mallow's Cp on the one hand and a stepwise selection procedure on the other hand (Myers, 1990).

Influence statistics DFFITS and DFBETAS (Myers, 1990) were used for detecting potential outliers and influential observations. Thus, from the normative sample of healthy controls subjects below twenty years of age or with fewer than eight years of education have been excluded as well as two more overly influential subjects.

The above procedure produced regression formulae for the various subtests which can be summarized as

$$Z = [f(T) - (b0 + b1*SEX + b2*AGE + b3*AGE2 + b4*SCHOOL + b5*SCHOOL2)] / c$$

where Z denotes the resulting z-score, f(T) the transformation of the raw subtest value T, SEX equals 0 for male and 1 for female patients, AGE denotes age (in years), AGE2 the square of the sex specifically centred age (i.e. (AGE-60)2 for males and (AGE-59)2 for females), SCHOOL the years of education and SCHOOL2 the square of the sex specifically centred years of education (i.e. (SCHOOL-13)2 for males and (SCHOOL-11)2 for females). The transformation f, the regression coefficients b1 to b5 as well as the standardization parameter c are given for each subtest in table 1.

Table 1: Summary of the variables and coefficients used in the regression model based calculation of z-scores in a healthy control sample.

Neuropsychological tests	n	f(T)	b_0	b_1	b_2	b_3	b_4	b_5	c	R^2	Adj. R^2
Psycho Motor speed											
Symbols counting (C.I.)	216	Log(T)	2.261		0,0089				0,231	0,183	0,179
			0.080		0,0013						
TMTA	332	Log(T)	2.714		0,0155	0,0002	-0,0173		0,307	0,362	0,357
			0,112		0,0014	0,0001	0,0046				
Attention											
AKT	427	Log(T)	2.907		0.0114	0.0001	-0.0216		0.236	0,379	0.375
			0.073		0.0009	0.0000	0,0031				
Digit-Symbol -Test	434	Sqrt(T)	8.357	0.240	-0,0456	-0.0006	0,0987	-0.0056	0,756	0.482	0,476
			0.247	0,076	0,0029	0.0001	0.0122	0,0024			
TMTB	208	Log(T)	3.676	0.087	0,0119		-0,0124		0,330	0,215	0,204
			0,160	0,048	0,0019		0,0065				
Language											
Semantic verbal fluency total	155	Log(T)	4.249		-0,0039		0,0081		0,212	0,074	0,062
			0.112		0.0015		0.0044				
Animals	351	Log(T)	3.325		-0,0053	-0.0001	0,0134		0,253	0,118	0,110
			0.090		0.0011	0.0000	0,0037				
Supermarket items	345	Log(T)	3.548		-0.0057	-0.0001			0,232	0,091	0,086
			0.068		0.0010	0.0000					
Tools	155	Log(T)	3.013	-0.178	-0.0075				0,360	0,101	0,089
			0,161	0,059	0,0025						
Phonematic verbal fluency	279	Sqrt(T)	5.217				0,0541		0,960	0,048	0,045
			0,193				0.0145				
Letter b	174	Sqrt(T)	3.099				0.0324		0,641	0,044	0,038
			0,155				0.0116				
Letter f	174	Sqrt(T)	2.885				0,0302		0,654	0,037	0,031
			0,158				0.0118				
Letter l	174	Sqrt(T)	3.098				0,0323		0,583	0,052	0,047
			0,141				0.0105				
BNT	253	Arcsin(Sqrt(T/15))	1,757	-0,096	-0,0050				0,165	0,168	0,162
			0,058	0,021	0,0009						
Memory											
Verbal memory immediate recall (VSRT)	363	T	10,939	0,710	-0,0679	-0,0018	0,0907		1,945	0,208	0,199
			0,677	0,215	0,0083	0,0005	0,0280				
Verbal memory total recall (VSRT)	363	T^c	4312,784	477.961	-38.9510	-0.8635	51.1425		795,397	0.340	0.333
			276.974	87.749	3.3908	0.2092	11.4431				
Verbal memory delayed recall (VSRT)	301	T	13,789	1,400	-0,0675	-0,0010	0,0579		2,125	0,200	0,189
			0,806	0,256	0,0101	0,0006	0,0331				
Executive Functions											
Nonverbal Fluency (Five Point Test)	144	Sqrt(T)	6.660		-0,0250		0,0316		0,767	0,167	0,155
			0,420		0,0054		0,0165				
Stroop (NAI)	124	Log(T)	3.338		0,0103		-0,0175		0,242	0,280	0,268
			0,138		0,0018		0,0055				
Planning (Maze Test -NAI)	122	Log(T)	3.083		0,0104		-0,0238		0,374	0,166	0,152
			0,213		0,0028		0,0084				
Interference (TMTA – TMTB)	207	Log(T)	2.876	0.154	0,0115				0,495	0,098	0,089
			0,176	0,070	0,0028						
Interference Test (C.I.)	214	Log(T)	2.825		0.0075		-0,0230	0,0022	0,227	0,228	0,217
			0,103		0,0013		0,0050	0,0011			

Number *n* of healthy controls, transformation *f*, regression coefficients b_1 *to* b_5 (with standard errors below) and standardization parameter *c* for each subtest.
Empty cells indicate that the corresponding adjustment variable is not used in the normalization of the respective subtest (No interaction of any two effects were selected).
In addition, R^2 and Adjusted R^2 are given for each subtest regression model.

In order to validate, the regression formulae z-scores for all controls in the present sample were calculated and showed distributions that can be expected from normally distributed variables.

Convergent validity was investigated calculating group differences on continuous variables using separate Wilcoxon rank-sum-tests for each cognitive variable with cognitive status (dementia vs. controls) as independent variable. Separate t-tests for each corresponding z-score with cognitive status (dementia vs. controls) as independent variable were also performed. Multiple tests were corrected by the method of Bonferroni-Holm and uncorrected p-values are

given with an indication of significance after correction (Hochberg & Tamhane, 1987).

In order to study the diagnostic utility of each test to predict dementia, we performed analyses for sensitivity, specificity, positive and negative predicted value, percent of correctly predicted patients and receiver operating characteristic (ROC) curve analysis. The area under the ROC curve (AUC) was used to measure the accuracy of discrimination between dementia patients and controls. Thus, the AUC allowed evaluation of how well each measure discriminated between AD patients and controls over the entire range of cut off points. Furthermore, a cut-off value was chosen for each subtest so that sensitivity and specificity are as close as possible. Sensitivity, specificity, positive and negative predicted value, percent of correctly predicted at the chosen cut-off value, as well as the AUC are given with 95% confidence intervals.

The reported p-values are the results of two-sided tests. P-values ≤ 0.05 were considered to be statistically significant. All computations have been performed using SAS software Version 9.1 (SAS Institute Inc., Cary, NC, USA, 2001).

3. Results

Separate Wilcoxon tests for each cognitive variable and separate t-tests for each corresponding z-score detected statistical group differences after correcting for multiple testing between dementia patients and controls. Values of all considered neuropsychological tests (cognitive variables) are listed as median and quartiles for raw scores and as mean and standard deviation for z scores in table 2.

Table 2: Medians, 25th and 75th percentile for raw scores and corresponding mean z scores with standard deviations for each neuropsychological variable for controls and AD patients

Neuropsychological tests	Controls (N=58)		AD patients (N=78)	
	Raw scores[a]	z Scores[b]	Raw Scores	z Scores
Psycho motor speed				
symbol counting (C.I.)	18.0 (14; 21)	-0.10 ± 1.22	27.0 (21; 33)	-1.83 ± 1.73
TMT-A	37.0 (29; 51)	-0.25 ± 1.01	62.0 (52; 83)	-1.56 ± 1.37
Attention				
AKT	31.5 (25; 39)	-0.06 ± 1.05	60.0 (45; 82)	-2.33 ± 1.97
Digit-symbol	40.0 (30; 50.5)	-0.27 ± 1.01	23.0 (15.5; 27.5)	-1.64 ± 1.45
TMT-B	92.5 (65; 116)	-0.45 ± 1.07	263.0 (146; 300)	-2.76 ± 1.57
Language				
Semantic fluency total	55.0 (48; 63)	-0.50 ± 1.46	35.0 (28; 42)	-2.68 ± 1.92
animals	21.0 (17; 26)	-0.33 ± 1.37	12.5 (10; 14)	-2.25 ± 1.49
supermarket items	22.0 (18; 26.5)	-0.47 ± 1.50	13.5 (10.5; 17)	-2.35 ± 1.55
tools	10.0 (8; 14)	-0.11 ± 0.99	7.0 (4; 9)	-1.35 ± 1.37
Phonematic fluency	31.0 (24; 41)	-0.28 ± 1.09	17.5 (11; 25)	-1.64 ± 1.01
letter b	10.0 (7; 14)	-0.50 ± 1.20	5.5 (4; 8)	-1.71 ± 1.04
letter f	9.5 (7; 13)	-0.16 ± 1.00	6.0 (3; 8)	-1.37 ± 1.13
letter l	11.0 (8; 15)	-0.23 ± 1.10	6.0 (4; 10)	-1.59 ± 1.18
Confrontation Naming (BNT)	14.0 (13; 14)	-0.40 ± 1.08	12.0 (10; 13)	-1.49 ± 1.23
Memory				
Selective Reminding immediate recall (VSRT)	8.0 (7; 9)	0.27 ± 0.87	4.0 (2; 5)	-1.46 ± 1.08
Selective Reminding total recall (VSRT)	49.0 (43; 54)	-0.08 ± 0.78	22.0 (18; 29)	-1.69 ± 0.71
Selective Reminding delayed recall (VSRT)	10.5 (9; 12)	-0.09 ± 0.82	3.0 (1;5)	-3.17 ± 1.08
Executive Functions				
Nonverbal Fluency (Five Point Test)	28.0 (21; 36)	-0.19 ± 1.16	14.0 (11; 16)	-1.81 ± 0.88
Stroop Test (NAI)	47.5 (42; 60.5)	-0.27 ± 1.00	72.0 (54; 100)	-1.81 ± 1.56
Planning (Maze Test -NAI)	35.0 (28.5; 44)	-0.20 ± 1.04	74.0 (52; 108)	-1.86 ± 1.61
Interference (TMTA – TMTB)	49.0 (34; 83)	-0.32 ± 1.29	210.0 (93; 226)	-2.36 ± 1.32
Interference Test (C.I.)	25.0 (21; 29)	-0.47 ± 1.02	35.0 (30; 43)	-1.93 ± 1.47

[a] Significant group effects ($p<0.001$) for all neuropsychological tests (raw scores) based on wilcoxon tests
[b] Significant group effects ($p<0.001$) for all neuropsychological tests (z- scores) based on t tests

Compared with the normative sample, the results of the dementia sample indicate that the largest decline from normal was observed on Verbal Selective Reminding Test delayed recall (-3.2 SD). Beyond this memory test, the greatest impairment was seen in the domain of attention (TMT-B and AKT), the domain of language (semantic verbal fluency) and executive function (difference score of TMT-A and TMT-B).

Table 3 presents the results of analyses for sensitivity, specificity, positive and negative predicted value, percent of correctly predicted at the chosen cut-off value for the raw score, as well as the AUC with 95% confidence intervals. Z scores produced similar results. However, for most cognitive variables the AUC was approximately five percent lower (but far from statistically significantly lower) which might be explained by the fact that the group differences with respect to age and education (both of which give a clear disadvantage to the demented group) are implicitly accounted for when using z-scores.

Table 3: Results of analyses of sensitivity, specificity, positive and negative predicted value, percent of correctly predicted at the chosen cut-off value as well as the receiver operating characteristics (AUC) with 95% confidence intervals.

Neuropsychological tests	Cut off Score	Sensitivity	Specificity	Positive Predicted Value	Negative Predicted Value	% correct	AUC
Psycho motor speed							
Symbol counting (C.I.) (dementia N=54; control N=55)	22	0.74 (0.60-0.85)	0.76 (0.63-0.87)	0.75 (0.62-0.86)	0.75 (0.62-0.86)	75.2 (66.0-83.0)	0.85 (0.78-0.93)
TMT-A (dementia N=61; control N=57)	52	0.77 (0.65-0.86)	0.75 (0.62-0.86)	0.77 (0.65-0.87)	0.75 (0.62-0.86)	76.3 (67.5-83.6)	0.86 (0.78-0.93)
Attention							
AKT (dementia N=77; control N=58)	42	0.83 (0.73-0.91)	0.81 (0.69-0.90)	0.85 (0.75-0.92)	0.78 (0.66-0.88)	82.2 (74.7-88.3)	0.90 (0.85-0.95)
Digit-Symbol-Test (dementia N=48; controls N=56)	29	0.81 (0.67-0.91)	0.79 (0.66-0.88)	0.76 (0.63-0.87)	0.83 (0.71-0.92)	79.8 (70.8-87.0)	0.88 (0.82-0.95)
TMT-B (dementia N=7; control N=54)	139	0.86 (0.42-0.99)	0.85 (0.73-0.93)	0.43 (0.18-0.71)	0.98 (0.89-0.99)	85.2 (73.8-93.0)	0.94 (0.87-1.00)
Language							
Semantic fluency total (dementia N=31; controls N=51)	43	0.81 (0.63-0.93)	0.80 (0.67-0.90)	0.71 (0.54-0.85)	0.87 (0.74-0.95)	80.5 (70.3-88.4)	0.87 (0.80-0.95)
Animals (dementia N=74; controls N=58)	15	0.78 (0.67-0.87)	0.78 (0.65-0.87)	0.82 (0.71-0.90)	0.74 (0.61-0.84)	78.0 (70.0-84.8)	0.87 (0.81-0.94)
Supermarket items (dementia N=36; controls N=52)	17	0.81 (0.64-0.92)	0.65 (0.65-0.89)	0.73 (0.56-0.85)	0.85 (0.72-0.84)	79.5 (69.6-87.4)	0.87 (0.80-0.94)
Tools (dementia N=31; controls N=51)	8	0.68 (0.49-0.83)	0.75 (0.60-0.86)	0.62 (0.44-0.78)	0.79 (0.65-0.89)	72.9 (60.9-81.3)	0.78 (0.68-0.88)
Phonematic fluency total (dementia N=32; controls N=49)	24	0.69 (0.50-0.84)	0.71 (0.57-0.83)	0.61 (0.43-0.77)	0.78 (0.63-0.89)	70.4 (59.2-80.0)	0.82 (0.72-0.91)
letter b (dementia N=30; controls N=51)	7	0.70 (0.51-0.85)	0.69 (0.54-0.81)	0.57 (0.39-0.73)	0.80 (0.65-0.90)	69.1 (57.9-78.9)	0.78 (0.69-0.88)
letter f (dementia N=30; controls N=50)	7	0.70 (0.51-0.85)	0.68 (0.53-0.80)	0.57 (0.39-0.73)	0.79 (0.64-0.90)	68.8 (57.4-78.6)	0.79 (0.69-0.88)
letter l (dementia N=31; controls N=49)	9	0.71 (0.52-0.86)	0.67 (0.52-0.80)	0.58 (0.41-0.74)	0.79 (0.63-0.89)	68.8 (57.4-78.6)	0.80 (0.71-0.90)
BNT (dementia N=70; controls N=57)	12	0.64 (0.52-0.75)	0.82 (0.70-0.91)	0.82 (0.69-0.91)	0.65 (0.53-0.76)	72.4 (63.8-80.0)	0.79 (0.71-0.86)
Memory							
Selective Reminding immediate recall (dementia N=78; controls N=58)	5	0.79 (0.69-0.88)	0.93 (0.83-0.98)	0.94 (0.85-0.98)	0.77 (0.66-0.86)	85.3 (78.2-90.8)	0.93 (0.89-0.97)
Selective Reminding total recall (dementia N=78; controls N=58)	34	0.91 (0.82-0.96)	0.93 (0.83-0.98)	0.94 (0.86-0.99)	0.88 (0.78-0.95)	91.9 (85.9-95.9)	0.98 (0.96-1.00)
Selective Reminding delayed recall (dementia N=78; controls N=58)	6	0.94 (0.86-0.98)	0.91 (0.81-0.97)	0.94 (0.86-0.98)	0.91 (0.81-0.97)	92.6 (86.9-96.4)	0.99 (0.97-1.00)
Executive Functions							
Five Point Test (dementia N=25; controls N=51)	18	0.80 (0.59-0.93)	0.80 (0.67-0.90)	0.67 (0.47-0.83)	0.89 (0.76-0.96)	80.3 (69.5-88.5)	0.90 (0.83-0.97)
Stroop Test (NAI) (dementia N=19; controls N=52)	60	0.74 (0.49-0.91)	0.73 (0.59-0.84)	0.50 (0.31-0.69)	0.88 (0.75-0.96)	73.2 (61.4-83.0)	0.83 (0.73-0.93)
(Maze Test (NAI) (dementia N=22; controls N=52)	50	0.82 (0.60-0.94)	0.83 (0.70-0.92)	0.67 (0.46-0.83)	0.91 (0.80-0.98)	82.4 (71.8-90.3)	0.85 (0.75-0.95)
Interference (TMTB-TMTA) (dementia N=7; controls N=50)	139	0.86 (0.42-0.99)	0.85 (0.73-0.93)	0.43 (0.18-0.71)	0.98 (0.89-0.99)	85.2 (73.8-93.0)	0.94 (0.87-1.00)
Interference (C.I.) (dementia N=51 controls N=55)	30	0.84 (0.71-0.93)	0.78 (0.65-0.88)	0.78 (0.65-0.88)	0.84 (0.71-0.93)	81.1 (72.3-88.0)	0.86 (0.78-0.93)

4. Discussion

The purpose of this study was to report standardization procedures, provide norms and evaluate the diagnostic utility of the Vienna Neuropsychological Test Battery (VNTB) in detecting Alzheimer`s Disease.

In a first step, the standardization of the test battery using a healthy sample of elderly was carried out. As reported in the literature we also found age, gender and schooling effects in our sample for most of the cognitive variables (Berres et al., 2000; Welsh et al., 1994). On the basis of multiple regression analyses, we calculated regressions based norms in order to provide z-scores for each single subtest. A computer based calculating program was developed to calculate corresponding z –scores for each cognitive variable.

Using the raw data and the regression based z-score, convergent validity analyses were performed separately for both data sets. We found very good discriminating power for the neuropsychological test battery in detecting AD dementia for data the set using raw scores and the data set using z scores, respectively. All single variables significantly separated control patients without cognitive impairment and AD patients on a group basis. Specifically, comparing the predictive accuracy of single neuropsychological tests using results from receiver operating characteristics analyses (area under the curve; AUC), we found that verbal memory testing using the selective reminding total recall and the delayed recall paradigm had very high discrimination accuracy. The AUC for total recall was 0.98 and for delayed recall was 0.99 with corresponding sensitivities ranging from 0.91 to 0.94 and specificities ranging from 0.91 to 0.93 respectively. Other psychometric measures including the AKT, TMT-B, Five Point Test and the difference score of the TMT test had also very good discriminative power with an area under the curve of greater than 0.90. This result is in accordance with recent studies investigating the discriminative ability of several cognitive tests in predicting AD patients (Chen et al., 2000; Grober et al., 2000). Taken together with prior results, this corroborates the hypothesis that verbal memory is one of the first cognitive functions being affected by the disease process and a good indicator as to whether AD dementia is present or not. Our data further demonstrate that performance on tests of attention and executive functions also have diagnostic value.

We also compared the pattern of cognitive loss of our dementia sample to another recent study which used z-scores for comparison purposes in a sample of patients (Grundman et al., 2004). In comparing z-scores across cognitive variables, we found that delayed recall memory performance, attention/executive functions and semantic verbal fluency showed the largest decline from normal. Grundman et al. (2004) also found verbal memory, verbal fluency and attention primarily impaired in their sample.

The limitations of the present study should be addressed. First, our healthy control sample consisted of patients of the General Hospital of Vienna, and as such they were not drawn from the general elderly population. However, careful medical examination and psychometric screening guaranteed that patients with cognitive problems were not included in the healthy control sample. Second, cognitive variables resulting in test results without normal distribution (e.g. recognition memory results of the VSRT) are not included in the study because regression based analysis was not possible due to the skewed distribution of the data.

We conclude that the newly developed VNTB can help discriminate patients with AD dementia from patients without dementia in patients reporting memory problems and seeking medical help. Thus, our study contributes to the knowledge regarding the diagnostic value of psychometric testing in patients complaining of memory problems in the setting of a memory clinic. The VNTB may also be of value in examining pre-dementia states because parts of the VNTB, in particular VSRT, AKT, TMT A and TMT B, AKT, digit-symbol test, have been shown to possess prognostic value in predicting whether a patient reporting memory problems converts to AD dementia within two years in a prior study (Lehrner et al., 2005).

5. Acknowledgements

We would like to thank Arinya Eller for proof-reading the manuscript. An EXCEL based copy of the statistical program for calculating z-scores can be obtained from gernot.petry@chello.at at a minimal charge. Test material can be obtained from the homepage of the Austrian Neuropsychological Society www.gnpoe.at/arbeitsgruppen/gerontoneuropsychologie.

References

BERRES, B., MONSCH, A. U., BERNASCONI, F., THALMANN, B. & STAHELIN, H. (2000). Normal ranges of neuropsychological tests for the diagnosis of Alzheimer's disease. Studies Health Technology and Information, 77, 195-195.

CHANDLER, M. J., LACRITZ, L. H., HYNAN, L. S., BARNARD, H. D., ALLEN, G. & DESCHNER, M. (2005). A total score for the CERAD neuropsychological battery. Neurology, 65, 102-106.

CHEN, P., RATCLIFF, G., BELLE, S. H., CAULEY, J. A., DEKOSKY, S. T. & GANEULI, M. (2000). Cognitive tests that best discrimiante between presymptomatic ad and those who remain nondemented. Neurology, 55, 1847-1853.

CROOK, T., SALAMA, M. & GOBERT, J. (1986). A computerized test battery for detecting and assessing memory disorders. In A. Bes, J. Cohn, S. Hoyer, J. P. Marc-Vergenes and H. M. Wisniewski (Eds.), Senile dementias: Early detection (pp. 79-85). London-Paris: John Libbey Eurotext.

FOLSTEIN, M., FOLSTEIN, S. & MCHUGH, P.R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

GATTERER, G. (1990). Alters-Konzentrations-Test (AKT). Göttingen: Hogrefe.

GOODGLASS, H. & KAPLAN, P. (1983). The assessment of aphasia and related disorders (2nd ed.). Philadelphia: Lea & Fabinger.

GROBER, E., LIPTON, R. B., HALL, C. & CRYSTAL, H. (2000). Memory impairment on free and cued selective reminding predicts dementia. Neurology, 54, 827-832.

GRUNDMAN, M., PETERSEN, R. C. & FERRIS, S. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging. Archives of Neurology, 61, 59-66.

HOCHBERG, Y. & TAMHANE, A. C. (1987). Multiple comparison procedures. New York: John Wiley.

IVNIK, R., MALEC, J. F. & SMITH, G. D. (1992). Mayo's older Americans normative studies: WAIS-R, WMS-R and AVLT norms for ages 56 through 97. Clinical Neuropsychology, 6, 1-104.

LEHRL, S. & FISCHER, B. (1997). Kurztest für cerebale Insuffizienz (c.I.-test). Ebersberg: Vless.

LEHRNER, J., GLEISS, A., MALY, J., AUFF, E. & DAL-BIANCO, P. (2006). Der Verbale Selektive Reminding Test (VSRT). Ein Verfahren zur Überprüfung Verbaler Gedächtnisfunktionen. Neuropsychiatrie, 20, 204-214.

LEHRNER, J., GUFLER, R., GUTTMANN, G., MALY, J., GLEISS, A. & AUFF, E. (2005). Annual conversion to Alzheimer's disease among patients with memory complaints attending an outpatient memory clinic: The influence of amnestic mild cognitive impairment and the predictive value of neuropsychological testing. Wiener Klinische Wochenschrift, 117/18, 629-635.

LEZAK, M. D. (1995). Neuropsychological assessment. New York: Oxford University Press.

LUIS, C. A., LOEWENSTEIN, D. A., ACEVEDO, A., BARKER, W. W. & DUARA, R. (2003). Mild cognitive impairment: Directions for future research. Neurology, 61, 438-444.

MCKHANN, G., DRACHMAN, D., FOLSTEIN, M., KATZMAN, R., PRICE, D. & STADLAN, E. M. (1984). Clinical diagnosis of Alzheimer disease. Neurology, 34, 939-944.

MEULEN, E. F. J., SCHMAND, B., VAN CAMPEN, J. P., DE KONING, S. J., PONDS, R. W. & SCHELTENS, P. (2004). The Seven Minute Screen: A neurocognitive screening test highly sensitive to various types of dementia. Journal of Neurology, Neurosurgery, and Psychiatry, 75, 700-705.

MORRIS, J. C., HEYMAN, A., MOHS, R. C., HUGHES, J. P., VAN BELLE, G. & FILLENBAUM, G. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology, 39, 1159-1165.

MYERS, R. H. (1990). Classical and modern regression with applications. Kent: PWS.

OSWALD, W. D. & FLEISCHMANN, U. M. (1997). Das Nürnberger-Alters-Inventar. Göttingen: Hogrefe.

PETERSEN, R. C. (2004). Mild cognitive impairment as a diagnostic entity. Journal of Internal Medicine, 256, 183-194.

POWLISHTA, K. K., VON DRAS, D. D., STANFORD, A., CARR, D. B., TSERING, C. & MILLER, J. P. (2002). The clock drawing test is a poor screen for mild dementia. Neurology, 59, 898-903.

REGARD, M., STRAUSS, E. & KNAPP, P. (1982). Children's production on verbal and non-verbal fluency tasks. Perceptual and Motor Skills, 55, 839-844.

REITAN, R. (1979). Trail Making Test (TMT). Tucson: Reitan Neuropsychology Laboratory.

STEINER, M. (1998). Das alternde Gedächtnis: Grenzen und Perspektiven – die MAC-Gedächtnisbatterie. Universität Wien: Wien.

TEWES, U. (1994). Hamburg-Wechsler-Intelligenztest für Erwachsene – Revision 1991 (HAWIE-R). Bern: Verlag Hans Huber.

WELSH, K., BUTTERS, N., MOHS, R., BEEKLY, D., EDLAND, S. & FILLENBAUM, G. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological test battery. Neurology, 44, 609-614.

YOUNGJOHN, J. R., LARRABEE, G. J. & CROOK, T. H. (1991). First-last names and the grocery list selective reminding test: Two computerized measures of everyday verbal learning. Archives of Clinical Neuropsychology, 6, 287-300.

Autoren

Mag. Dr. Johann Lehrner
Klinischer Psychologe und Gesundheitspsychologe
Abteilung für Neuropsychologie

Universitätsklinik für Neurologie
Medizinische Universität Wien
Währingergürtel 18-20
A-1097 Wien
Johann.Lehrner@meduniwien.ac.at

Ass. Prof. Dr. Joachim Maly
Klinischer Psychologe und Gesundheitspsychologe
Abteilung für Neuropsychologie

Universitätsklinik für Neurologie
Medizinische Universität Wien
Währingergürtel 18-20
A-1097 Wien
Joachim.Maly@meduniwien.ac.at

Dr. Andreas Gleiß
Besondere Einrichtung für Medizinische Statistik und Informatik

Medizinische Universität Wien
Währingergürtel 18-20
A-1097 Wien
Andreas.Gleiss@meduniwien.ac.at

o.Univ. Prof. Dr. Eduard Auff
Facharzt für Neurologie und Psychiatrie
Vorstand der Universitätsklinik für Neurologie

Medizinische Universität Wien
Währingergürtel 18-20
A-1097 Wien
Eduard.Auff@meduniwien.ac.at

ao. Univ. Prof. Dr. Peter Dal-Bianco
Facharzt für Neurologie und Psychiatrie
Leiter der Gedächtnisambulanz

Universitätsklinik für Neurologie
Medizinische Universität Wien
Währingergürtel 18-20
A-1097 Wien
Peter.Dal-Bianco@meduniwien.ac.at