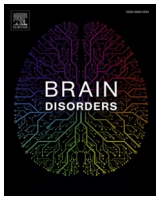
Brain Disorders 5 (2022) 100028

Contents lists available at ScienceDirect

Brain Disorders

journal homepage: www.sciencedirect.com/journal/brain-disorders



Neuropsychological prediction of dementia using the neuropsychological test battery Vienna – A retrospective study

Anna Garcia Rosas, Elisabeth Stögmann, Johann Lehrner *

Department of Neurology, Medical University of Vienna, Vienna, Austria

ARTICLE INFO

Keywords:
Cognitive functions
Neuropsychological testing
Subjective cognitive decline
Mild cognitive impairment
Alzheimer's disease

ABSTRACT

Before clinical cognitive deteriorations appear, the progression of dementia can be predicted by neuropsychological tests. The present study assessed the validity of the Neuropsychological Test Battery Vienna (NTBV) and the discrimination ability between diagnostic groups, subjective cognitive decline (SCD), mild cognitive impairment (MCI), Alzheimer's disease (AD) and healthy control (HC). The same patients were surveyed in a follow-up assessment after a mean interval of 25.96 months to examine cognitive performance and disease progression. Differences between NTBV subtests in diagnostic groups were found. The domain (verbal) memory identified best those persons who developed dementia over time. The results of the receiver operating characteristic (ROC) analysis showed that areas under the curves ranged from 0.20 to 0.91. The test-retest reliability for NTBV scores ranged from 0.41 to 0.85 in the HC group, 0.15 – 0.87 for the SCD group, 0.39 – 0.80 for the MCI group and 0.51– 0.82 in the total sample group. Psychometric criteria were shown to range from low values to excellent values. The domain memory achieved the best discrimination power for detecting dementia.

Introduction

Dementia is a disorder that occurs mainly in elderly people with different etiologies such as neurodegeneration, vascular causes, metabolic causes and many more [, 2]. The most common form is the slowly but steadily progressing and irreversible Alzheimer`s disease (AD) . The disorder is characterized by a cognitive decline from a previous level of functionality [2,4]. The typical development of dementia involves three phases. The first phase contains subjective cognitive decline (SCD), followed by mild cognitive impairment (MCI) and AD-dementia [5, 6]. Knowing about the disease at an early stage might help to delay cognitive deterioration by using preventive action such as cognitive, physical and social intervention strategies.

The most salient predictors of AD-dementia are verbal and visuospatial memory, which have been reported as being among the earliest signs related to increased risk of developing AD-dementia [7–11]. Word-finding problems may also be found [11]. Planning strategies and abstract thinking may worsen continuously. Everyday activities such as financial activities, driving or personal hygiene may be impaired [, 12, 13].

The Neuropsychological Test Battery Vienna (NTBV) [14] was developed to assess neurocognitive functions in elderly patients. The NTBV assesses a broad range of cognitive abilities commonly affected by SCD, MCI and AD. It consists of different subtests, which in turn can be subdivided into different domains. These domains are attention, language, executive functioning and memory.

Neuropsychological test batteries are less practical in a clinical setting due to longer administration times compared to screening measures such as the Mini Mental State Examination (MMSE). However, as these tests contain much more information, the NTBV variables may tell which aspects of cognition are overlooked or underweighted when the clinical diagnosis of MCI is made. For example, executive functioning is thought to be poorly evaluated in the MMSE. We focused on the diagnostic accuracy of different cognitive tests in different cognitive domains. As such, we were interested in which measures yielded a higher diagnostic accuracy when compared to other measures.

The aim of the study was to evaluate if the NTBV is a valid neuropsychological measure in diagnosing dementia. We hypothesized that specific neurocognitive tests are able to predict conversion to AD-dementia and that tests of the memory domain will be superior to other measures. Furthermore, we predicted differences between NTBV subtests across the first and second examinations and between NTBV subtests for converters to AD-dementia and non-converters. Likewise, we predicted differences for all subtests in diagnostic groups except for

* Correspondent author at: Department of Neurology Medical University of Vienna Währinger Gürtel 18-20 A-1090 Wien.
E-mail address: johann.lehrner@meduniwien.ac.at (J. Lehrner).

https://doi.org/10.1016/j.dscb.2021.100028
Received 17 June 2021; Received in revised form 13 December 2021; Accepted 13 December 2021
Available online 15 December 2021

the comparison between healthy controls (HC) and SCD groups. Additionally, we predicted excellent validity (high sensitivity and specificity) and excellent test-retest reliability for the NTBV.

Methods

Data from 358 participants was analyzed. This study is a prospective cohort study encompassing consecutive, community-dwelling patients complaining of cognitive problems who were examined in the memory outpatient clinic for assessment of possible cognitive impairments. The participants were examined at two time points. Twelve months after the first appointment, the participants received an invitation for a follow-up appointment that they could respond to within a time frame of 48 months. So, the participants were surveyed during two measurements over 12 – 48 months (mean time interval for HC was 29.3 ± 13.2 months, 29.1 ± 11.7 months for SCD, and 24.9 ± 10.7 months for MCI, respectively). This long time frame was not intentional. It occurred because patients did not respond in time and had to be reminded. For the purpose of the analysis, patients were divided into the subgroups HC, SCD, MCI and AD. SCD was diagnosed by the criteria of Jessen et al. [15]. MCI was determined by the guidelines of Petersen [16] and to assess healthy functioning in HC. AD was established by the criteria of DSM-4 [4] and NINCDS-ADRDA [17] in a consensus conference with neuropsychologists, neurologists and other collaborators involved in the study of the cognitive status of the subjects. Only patients who performed the extended version of the NTBV at the first examination were considered in this study. The participants had no history of cerebral vascular pathology or severe head injury and were not diagnosed with AD-dementia at the first examination. In addition, they had no current psychiatric diagnosis and no medical condition that would lead to severe cognitive impairments. The age of the participants was 51 years or older. The criteria were evaluated by a diagnostic interview.

Neuropsychological measures

To assess premorbid IQ, the Wortschatz test (WST-IQ) [18], a standardized vocabulary test, was used. For measuring depression, the patients filled in the Beck Depression Inventory (BDI-II) [19]. Cognitive screening was achieved via the MMSE [20].

The extended version of the NTBV was applied to all participants who achieved an MMSE score over 23. The neuropsychological assessment with the NTBV for each participant lasted approximately 60 min and was performed within one test session.

Attention

The domain attention was assessed by the Alters-Konzentrations-Test (AKT), a geriatric cancelation test [21], the digit-symbol subtest of the German Wechsler Adult Intelligence Scale-Revised [22], the symbol counting task from the cerebral insufficiency test (C.I.) [23], the Trail Making Test B (TMT B) [24] and the score difference of the Trail Making Test A (TMT A) and TMT B. To complete the task of the AKT, the participant is instructed to cross a determined sign in a row with other similar signs. The time needed to complete the task and the achieved score of the AKT are used to calculate the test score. The digit-symbol test consists of numbers from one to nine, which are coded with symbols. Below the code is a row of numbers one to nine repeated several times in a random order. The task is to write as many symbols as possible under the corresponding number in 90 s. The test score is made up of the number of written symbols. To complete the C.I., the participant has to count all the squares represented with other symbols as quickly as possible within 60 s. The time needed is used for the test score. In the TMT-B, the participant has to connect letters and numbers in ascending and alternating order. The test score results from the time taken in seconds to complete the task.

Executive functioning

Executive functioning is assessed in the NTBV by the subtests TMT A [24], the Five-Point Test [25], the Planning Maze Test and the Stroop test from the Nürnberger Alters-Inventar [26], the interference test from the C.I. [23] and the phonematic verbal fluency test (PWT) [27]. The TMT A consists of digits that must be connected with each other in ascending order. The test score is composed of the time taken in seconds to complete the task. For the Five-Point Test, the participant receives a sheet with several squares including five dots. The participant is asked to draw as many different figures as possible by connecting the dots with straight lines within three minutes. The test score is the number of different figures drawn. To complete the maze test, the participant has to find a way out of a maze. The time needed to get from the start to the endpoint is assessed, as well as the number of errors. The total score is calculated as the quotient of the difference between the maximum number of errors and the errors actually made and the time needed. The Stroop test consists of two tasks such as naming colors and naming words written in different colors. The time needed, the correct naming and the difference of both tests are used to calculate the final score. For the interference test, a row with several letters of A and B in different orders is shown to the participant. The participant is instructed to name the letters in reverse order as quickly as possible. The test score is calculated from the total score of correct naming and the time needed to perform the test. The PWT includes three tasks in which the participant has to name as many words as possible starting with the letters b, f and l. Each task last 60 s. The test score is obtained by counting the named words.

Language

The Boston Naming Test (BNT) [28] and the semantic verbal fluency test (SWT) [27] are used to examine language. The BNT total score is derived from the number of correctly named images. For the SWT, the participant is asked to name as many animals, grocery items and tools as possible. Each task lasts 60 s. The sum of all named words forms the test score.

Memory

The Verbal Selective Reminding Test (VSRT) [29], with the subtests of immediate recall, total recall, delayed recall and recognition, is used to assess memory [30, 31]. The task of the VSRT is to recall 15 listed grocery items that are shown visually to the participant in five trials. The score for VSRT immediate recall is derived from the number of correct answers in the first trial. The score for learning performance is calculated by adding the number of correct answers from the first to the fifth trial. After 20 min, the participant has to recall all 15 words. The score for delayed recall results from the number of correct answers. The final task is a recognition using old vs new grocery items.

Psychometric properties of the NTBV

Cronbach's alpha (α) for the NTBV was conducted for different diagnostic groups with values ranging from 0.87 to 0.89 as well for the total sample with a high internal consistency (0.83 – 0.93) [14, 32, 33]. Consistency over time was measured in recent studies with Pearson's correlations ranging from 0.69 to 0.94 [14, 32, 33]. Objectivity can be largely ensured by the detailed description of the instructions and the evaluation of the NTBV test scores [34]. Construct validity was determined [35] using principal component analysis and subsequent orthogonal Kaiser-Varimax rotation in six iterations. Six factors with intrinsic values higher than 1 were considered; these represented 65.3% of the total variance. Furthermore, a variable clustering procedure based on 250 cognitively healthy subjects resulted in a six-cluster solution of the NTBV [13].

Sample characteristics

A total of 358 patients ranging from 51 to 92 years (*Mdn* =69, *SD* = 9.12) were included in the study. The mean interval between the first examination and the follow-up examination was 25.96 months (*SD* = 11.28). 175 males (48.9%) and 183 females (51.1%) were included. The duration of formal education was between six and 24 years (*Mdn* = 11.00, *SD* = 4.12). Table 1 shows the baseline characteristics of the total sample and the different diagnostic groups.

At the baseline examination, 270 patients were diagnosed with MCI, 46 as HC and 42 with SCD. At follow-up, 48 patients converted to AD from the total of 270 MCI patients. 183 patients were still diagnosed as MCI and 39 were diagnosed as SCD. All 46 HC remained stable. Of the 42 SCD, 17 converted to MCI. The other 25 patients were still categorized as SCD.

Statistical analyses

Statistical analyses were conducted using IBM SPSS Statistics for Windows, Version 25.0. In order to control the type I error, Bonferroni correction was used. Because of the considerable difference of the sample size in diagnostic groups and a no-interval scaling of the dependent variable, nonparametric methods were applied for all statistical analyses. Demographic and clinical characteristics are described by their median due to a non-normal distribution. Cross-tabulation was performed to assess progression rates from the first to the second examination. To calculate the mean difference of the NTBV scores in all diagnostic groups, Kruskal–Wallis analysis was used. Post-hoc Dunn Bonferroni analyses were conducted to correct for multiple testing and to reveal rates of comparison of each diagnostic group. Next, one-tailed Friedman tests were performed for all diagnostic groups and for converters to AD and non-converters to AD to explore the difference between the two examinations. Converters to AD-dementia and non-converters were established. Further, the Mann–Whitney U test was performed to compare the NTBV scores and the demographic and clinical variables between converters and non-converters. An internal consistency check for all items of the NTBV and different diagnostic groups was calculated using Cronbach's alpha. In addition, test-retest reliability was assessed by Pearson's correlation with the whole sample and all subsamples within a time interval of 12 – 48 months. Test-retest reliability was not assessed for converters to AD-dementia vs non-converters. To assess the predictive validity of NTBV scores for differentiating diagnostic groups, receiver operating characteristic (ROC) curves including area under the curve analysis (AUC) with AD as a positive condition were used to determine sensitivity, specificity and cut-off scores. These were chosen by the Youden Index. An ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. An ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The true positive rate is also known as sensitivity. The false positive rate is also known as probability of false alarm and can be calculated as (1 − specificity). Next, stepwise logistic regression was conducted to assess more detailed information about the discriminating ability of the NTBV among converters to AD-dementia and non-converters.

Table 1
Demographic and clinical characteristics of the sample.

	HC (*N*= 46)	SCD (*N*= 42)	MCI (*N*= 270)	Total (*N*= 358)
Age	61 (51–76) **	66 (50–86) **	70 (50–92) **	69 (50–92) **
Male/ Female	30/16	21/21	124/146	175/183
Education	12 (6–24)	12 (8–22)	11 (8–22)	11 (6–24)
BDI-II	2 (0–29)	8 (0–31)	9 (0–50)	8 (0–50)
MMSE	29 (27–30) **	29 (26–30) **	28 (22–30) **	28 (22–30) **
WST-IQ	116 (90–139)	118 (88–133)	110 (77–140)	110 (77–140)

Note. Median (Minimum - Maximum).
** $p \leq 0.001$.
HC Healthy controls, *SCD* Subjective cognitive decline, *MCI* Mild cognitive impairment, *BDI-II* Beck-Depression-Inventory, *MMSE* Mini Mental State Examination, *WST-IQ* Wortschatztest Intelligenzquotient.

Results

Kruskal–Wallis analysis showed significant differences for all NTBV subtests ($p \leq 0.01$) using diagnosis as a grouping variable at the first and second examinations. Detailed information is shown in Tables 2a and 2b. Post-hoc Dunn Bonferroni analyses were conducted to correct for multiple testing and to reveal rates of comparison of each diagnostic group. At the first examination, no significant differences between the HC and SCD groups could be shown except for SWT. Looking at the second examination, there were no significant differences between diagnostic groups HC and SCD except for the NTBV scores Planning Maze Test ($p \leq 0.01$) and Planning Maze Test total/time ($p < 0.05$). The other group comparisons HC – MCI ($p < 0.05$), HC – AD ($p < 0.001$), SCD – MCI ($p < 0.05$) and SCD – AD ($p < 0.001$) confirmed significant mean differences of NTBV scores at both examinations. The remaining group comparison MCI – AD yielded significant differences ($p \leq 0.01$) except for the NTBV scores C.I. Symbols (p =0.33), PWT (p =0.19) and Stroop color words – color (p =0.11). One-tailed Friedman tests analysing NTBV subtest scores across the baseline and follow-up examinations for the HC, SCD and MCI diagnosis groups failed to show significant differences for all subtests. Mann–Whitney U tests were calculated to

Table 2a
NTBV scores using diagnosis as grouping variable at first examination.

	HC (*N*= 46)	SCD (*N*= 42)	MCI (*N*= 270)
Attention			
AKT	26.50 (11.00)	28.00 (13.00)	36.00 (15.00)
AKT TT	2.06 (0.87)	1.91 (0.90)	1.50 (0.59)
Digit-Symbol	50.50 (18.00)	47.50 (13.00)	38.00 (16.00)
TMT B	74.50 (48.00)	80.00 (36.00)	107.00 (54.00)
C.I. Symbols	18.00 (5.00)	18.50 (6.00)	21.00 (8.00)
TMT B – TMT A	38.50 (36.00)	43.50 (25.00)	62.00 (45.00)
Language			
SWT	76.50 (22.00)	57.50 (18.00)	50.00 (17.00)
BNT	15.00 (0.00)	14.00 (1.00)	14.00 (1.00)
Memory			
VSRT immediate	9.00 (2.00)	8.00 (3.00)	7.00 (3.00)
VSRT total	55.00 (13.00)	51.50 (12.00)	45.00 (15.00)
VSRT delayed	12.00 (3.00)	11.00 (4.00)	9.00 (4.00)
VSRT recognition	14.75 (0.63)	15.00 (1.00)	14.50 (1.50)
Executive function			
TMT A	32.50 (14.00)	36.50 (14.00)	44.00 (20.00)
PWT	40.50 (13.00)	35.00 (16.00)	28.00 (14.00)
5 Point	35.00 (11.00)	32.00 (11.00)	25.00 (14.00)
Stroop color	21.00 (7.00)	22.00 (6.00)	25.00 (7.00)
Stroop words	37.00 (10.00)	42.00 (13.00)	50.00 (19.00)
Stroop TT	0.84 (0.31)	0.84 (0.30)	0.69 (0.27)
Stroop difference	16.00 (7.00)	20.00 (12.00)	25.00 (15.00)
Planning Maze	29.50 (18.00)	35.50 (17.00)	40.00 (25.00)
Planning Maze TT	0.50 (0.41)	0.43 (0.23)	0.38 (0.24)
Interference C. I.	19.00 (6.00)	20.00 (5.00)	24.00 (9.00)
Interference C.I. TT	1.79 (0.58)	1.67 (0.41)	1.42 (0.50)

Note. Median (Interquartile range), significant $p < 0.01$.
HC Healthy controls, *SCD* Subjective cognitive decline, *MCI* Mild cognitive impairment,.
AKT Alters Konzentrations Test, *TT* total/time, *TMT B* Trail Making Test Version B,.
C.I. Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *SWT* Semantische.
Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *PW*.
Phonematische Wortflüssigkeit,.

Table 2b
NTBV scores using diagnosis as grouping variable at second examination.

	HC (*N*= 46)	SCD (*N*= 64)	MCI (*N*= 200)	AD (*N*= 48)
Attention				
AKT	27.50 (10.00)	30.50 (11.00)	36.00 (16.00)	50.00 (30.00)
AKT TT	2.08 (0.69)	1.76 (0.72)	1.50 (0.64)	1.08 (0.66)
Digit-Symbol	48.50 (17.00)	48.00 (13.00)	36.00 (16.00)	28.00 (12.00)
TMT B	69.50 (34.00)	79.00 (34.00)	118.00 (87.00)	225.00 (99.00)
C.I. Symbols	17.00 (5.00)	20.00 (7.00)	23.00 (9.00)	25.00 (15.00)
TMT B – TMT A	36.50 (28.00)	45.00 (32.00)	72.50 (68.00)	158.50 (77.00)
Language				
SWT	74.00 (25.00)	60.00 (11.00)	48.00 (18.00)	38.00 (9.00)
BNT	15.00 (0.00)	15.00 (1.00)	14.00 (2.00)	13.00 (2.00)
Memory				
VSRT immediate	9.00 (3.00)	8.00 (2.00)	6.00 (3.00)	4.50 (3.00)
VSRT total	55.50 (15.00)	52.00 (15.00)	41.00 (15.00)	26.00 (12.00)
VSRT delayed	12.00 (4.00)	11.00 (4.00)	8.00 (6.00)	2.50 (3.00)
VSRT recognition	15.00 (1.00)	15.00 (0.50)	14.00 (2.50)	9.00 (4.63)
Executive function				
TMT A	31.00 (16.00)	36.00 (17.00)	46.00 (22.00)	70.50 (36.00)
PWT	41.00 (13.00)	36.00 (13.00)	28.00 (15.00)	22.00 (12.00)
5 Point	35.50 (12.00)	32.00 (11.00)	26.00 (12.00)	18.00 (15.00)
Stroop color	21.50 (6.00)	22.00 (4.00)	26.00 (8.00)	28.00 (11.00)
Stroop words	39.50 (13.00)	44.00 (15.00)	52.00 (23.00)	75.00 (55.00)
Stroop TT	0.92 (0.33)	0.79 (0.31)	0.69 (0.32)	0.43 (0.34)
Stroop difference	17.00 (8.00)	21.00 (12.00)	25.00 (18.00)	50.00 (34.00)
Planning Maze	26.00 (12.00)	34.50 (19.00)	40.00 (28.00)	64.00 (40.00)
Planning Maze TT	0.58 (0.29)	0.44 (0.22)	0.36 (0.26)	0.21 (0.18)
Interference C. I.	20.00 (8.00)	20.00 (6.00)	25.00 (9.00)	33.50 (15.00)
Interference C.I. TT	1.70 (0.67)	1.70 (0.47)	1.36 (0.52)	1.00 (1.98)

Note. Median (Interquartile range), significant $p < 0.01$.
HC Healthy controls, *SCD* Subjective cognitive decline, *MCI* Mild cognitive impairment, *AD* Alzheimer's disease, *AKT* Alters Konzentrations Test, *TT* total/time, *TMT B* Trail Making Test Version B, *C.I.* Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *SWT* Semantische Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *PWT* Phonematische Wortflüssigkeit,.

compare NTBV subtest scores and other relevant variables between converters to AD and non-converters at the second examination. Significant differences were confirmed for all subtests ($p < 0.001$). Table 3 shows the median of test scores, years of education and age as well as the effect size.

Internal consistencies were computed for all NTBV scores and diagnostic groups using Cronbach's alpha at the follow-up examination. The diagnostic groups HC ($\alpha = 0.83$), SCD ($\alpha = 0.87$), MCI ($\alpha = 0.82$) and the total sample ($\alpha = 0.86$) revealed a high internal consistency ($\alpha > 0.80$). Cronbach's alpha for AD achieved a value of 0.79.

Test-retest reliability analysis for different diagnostic groups failed to reach an excellent test-retest reliability ($r > 0.80r$) except for the NTBV scores of digit-symbol (r= 0.81) and VSRT recognition (r= 0.87) in the SCD group and Stroop color words – color (r= 0.85) and Stroop color words – words (r= 0.81) in the HC group. Also, the digit-symbol subtest

Table 3
Mann-Whitney-U-Test between converters and non-converters based on median. Effect sizes are interpreted as small effect with r= 0.10, as medium effect with r= 0.30 and as large effect with r= 0.50.

	Non-converters	Converters	r
AKT	33.00 (12.69)	47.50 (18.24	0.33**
AKT total/time	1.62 (0.53)	1.13 (0.39)	0.33**
Digit-Symbol	43.00 (12.30)	29.00 (10.58)	0.31**
C.I. Symbols	20.00 (6.48)	24.50 (8.18)	0.24**
TMT A	40.00 (15.69)	58.00 (29.54)	0.29**
TMT B	96.00 (51.39)	172.50 (76.48)	0.37**
SWT	54.00 (15.36)	39.00 (13.15)	0.33**
PWT	31.00 (11.68)	24.50 (11.21)	0.17**
BNT	14.00 (1.21)	14.00 (1.89)	0.22**
VSRT immediate recall	8.00 (2.02	5.00 (1.63)	0.38**
VSRT total recall	48.00 (10.17)	31.00 (8.09)	0.48**
VSRT delayed recall	10.00 (2.96)	4.00 (2.81)	0.48**
VSRT recognition	14.50 (1.50)	11.50 (3.47)	0.43**
5 Point Test	29.00 (9.82)	18.00 (8.64)	0.32**
Stroop color words-color	24.00 (5.75)	29.50 (9.31)	0.29**
Stroop color words-words	46.00 (13.89)	70.50 (24.23)	0.35**
Stroop total/time	0.74 (0.22)	0.49 (0.21)	0.34**
Stroop color words difference	23.00 (11.65)	41.50 (18.12)	0.33**
Planning Maze Test	37.00 (21.24)	55.50 (25.64)	0.31**
Planning Maze Test total/time	0.41 (0.19)	0.26 (0.15)	0.30**
TMTB-TMTA difference	55.50 (44.79)	120.00 (60.94)	0.35**
Interference C. I. time	22.00 (6.68)	29.50 (10.37)	0.32**
Interference C.I. total/time	1.55 (0.43)	1.13 (0.34)	0.33**
MMSE	28.50 (1.33)	26.00 (1.83)	0.39**
BDI-II	8.00 (8.02)	6.50 (8.65)	0.08
WST IQ	110.00 (12.82)	110.00 (13.37)	0.02
Education	12.00 (4.18)	10.50 (3.63)	0.09
Gender (male/ female)	48.7/ 51.3	50.0/ 50.0	0.01
Age	67.00 (8.90)	75.00 (7.58)	0.30**

Note. Median ± interquartile range, Gender in percent, ** $p \leq 0.01$.
AKT Alters Konzentrations Test, *C.I.* Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *TMT B* Trail Making Test Version B, *SWT* Semantische Wortflüssigkeit, *PWT* Phonematische Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *MMSE* Mini Mental State Examination, *BDI-II* Beck-Depressions-Inventory, *WST IQ* Wortschatztest Intelligenzquotient.

in the total sample reached a high test-retest reliability (r= 0.82). Test-retest reliability for NTBV scores ranged from 0.41 to 0.85 in the HC group, 0.15 – 0.87 for the SCD group, 0.39 – 0.80 for the MCI group and 0.51 – 0.82 in the total sample group. See Table 4.

ROC curve analysis was performed to assess the predictive power of the NTBV subtests. The scores of the NTBV subtests at the second examination were used as predictors, and patients who converted to AD were used as a positive condition. Cut-off scores were chosen by the Youden Index. Based on this cut-off score, sensitivity, specificity, PPV, NPV, LR+, LR- and percentage of correctly predicted patients were calculated. Detailed information subtests are shown in Tables 5a and 5b.

The stepwise logistic regression model with NTBV subtests as predictor variables with converters to AD-dementia as a positive variable was significant ($p < 0.001$) with $X^2 = 142.66$ and R^2 (Nagelkerke) = 0.64 for the entire sample. The subtests VSRT delayed recall, VSRT total recall, VSRT recognition, TMT B, PWT and Stroop color words – words yielded a significant B. The effect of R^2 by Nagelkerke indicates a strong effect with a value higher than f= 0.40 (f= 1.33) [36]. The classification table of the stepwise logistic regression analysis shows 92.1% correctly predicted patients in the total sample. Thus, 97.0% of non-converters and 63.0% of converters were correctly classified with the logistic regression model.

Discussion

The aim of the presented study was to evaluate the psychometric properties and the predictive power of the NTBV for early prediction of dementia. For this purpose, reliability and discrimination accuracy of the NTBV scores were examined for determining diagnostic groups.

Table 4
Test-retest reliability coefficients for subgroups (Pearson).

	HC (*N*= 46)	SCD (*N*= 64)	MCI (*N*= 200)	Total sample (*N*= 358)
Attention				
AKT	0.59**	0.62**	0.60**	0.64**
AKT TT	0.68**	0.61**	0.66**	0.72**
Digit-Symbol	0.80**	0.81**	0.80**	0.82**
TMT B	0.72**	0.66**	0.63**	0.71**
C.I. Symbols	0.41**	0.15	0.50**	0.51**
TMT B - TMT A	0.67**	0.48**	0.56**	0.61**
Language				
SWT	0.76**	0.62**	0.66**	0.76**
BNT	0.61**	0.42**	0.60**	0.67**
Memory				
VSRT immediate	0.60**	0.31**	0.39**	0.52**
VSRT total	0.76**	0.67**	0.63**	0.78**
VSRT delayed	0.66**	0.53**	0.64**	0.73**
VSRT recognition	0.51**	0.87**	0.42**	0.58**
Executive function				
TMT A	0.61**	0.59**	0.55**	0.65**
PWT	0.69**	0.73**	0.69**	0.75**
5 Point	0.47**	0.57**	0.66**	0.67**
Stroop color	0.85**	0.67**	0.64**	0.72**
Stroop words	0.81**	0.73**	0.71**	0.75**
Stroop TT	0.67**	0.74**	0.72**	0.75**
Stroop difference	0.66**	0.51**	0.68**	0.69**
Planning Maze	0.65**	0.65**	0.51**	0.60**
Planning Maze TT	0.61**	0.58**	0.66**	0.67**
Interference C. I.	0.59**	0.73**	0.56**	0.64**
Interference C.I. TT	0.77**	0.67**	0.61**	0.70**

** $p \leq 0.01$, * $p \leq 0.05$; *HC* Healthy controls, *SCD* Subjective cognitive decline, *MCI* Mild cognitive impairment, *AKT* Alters Konzentrations Test, *TT* total/time, *TMT B* Trail Making Test Version B, *C.I.* Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *SWT* Semantische Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *PWT* Phonematische Wortflüssigkeit,.

40.5% of the SCD group converted to MCI and 17.8% of the MCI group converted to AD-dementia. This is similar to the results observed in prior investigations [7, 37, 38]. No patient progressed from SCD directly to AD-dementia. Only patients with MCI converted to AD at the second examination. This supports the assumption of the typical progression of dementia, starting from SCD to MCI and leading to AD-dementia over the course of time [5, 6]. Some MCI patients may not convert to AD-dementia during the study interval. However, they may convert at a later time point.

As expected, the HC and SCD groups did not show significant differences between NTBV scores apart from the subtest SWT at the first examination and the Planning Maze Test at the second examination. This underlines the criteria for SCD, showing normal age-, gender-, and education-adjusted performance on standardized cognitive tests [15, 39–41]. The MCI and AD groups did not differ in three subtests (C.I. Symbols, PWT, Stroop color words – color) at the baseline examination, while all other group comparisons showed significant differences for all subtests at both examinations. The sample included patients at very different stages of cognitive impairment. Some of them had cognitive deterioration in only one cognitive domain, while others showed deficits in various domains. Our findings underline the good discrimination power of the NTBV between diagnostic groups.

For the domain attention, all variables showed good predictive accuracy except for the subtest (TMT B – TMT A). Only one subtest of the language domain showed good discrimination power (SWT). Comparing the results to a prior study [14], differences can be noticed. Regarding the domain executive function, all subtests approached values over 0.70 for AUC except for the difference score of Stroop color words and PWT. These findings are consistent with prior study results [14].

Table 5a
Results of analyses of sensitivity, specificity, percent correctly predicted at the chosen cut-off value and receiver operating characteristics (AUC).

Predictor variable	Cut off	Sensitivity	Specificity	% correct	AUC
Attention					
AKT ($N_1 = 48$; $N_2 = 310$)	36.50	0.83	0.61	63.97	0.78
AKT TT ($N_1 = 48$; $N_2 = 310$)	1.38	0.75	0.71	71.51	0.78
Digit-Symbol ($N_1 = 48$; $N_2 = 310$)	36.50	0.75	0.67	68.16	0.76
TMT B ($N_1 = 48$; $N_2 = 310$)	113.50	0.83	0.67	69.27	0.81
C.I. Symbols ($N_1 = 48$; $N_2 = 309$)	19.63	0.83	0.50	54.34	0.70
TMT B – TMT A ($N_1 = 47$; $N_2 = 310$)	240.50	1.00	0.01	14.01	0.20
Language					
SWT ($N_1 = 48$; $N_2 = 310$)	42.50	0.63	0.81	78.49	0.78
BNT ($N_1 = 48$; $N_2 = 309$)	13.50	0.46	0.80	75.35	0.68
Memory					
VSRT immediate ($N_1 = 48$, $N_2 = 310$)	5.50	0.60	0.86	82.68	0.82
VSRT total ($N_1 = 48$; $N_2 = 310$)	41.50	0.94	0.71	74.02	0.91
VSRT delayed ($N_1 = 48$; $N_2 = 309$)	7.50	0.90	0.78	79.55	0.90
VSRT recognition ($N_1 = 48$; $N_2 = 309$)	13.25	0.73	0.85	83.66	0.85
Executive function					
TMT A ($N_1 = 48$; $N_2 = 310$)	53.50	0.58	0.79	76.26	0.75
PWT ($N_1 = 48$; $N_2 = 309$)	21.50	0.42	0.81	75.63	0.64
5 Point ($N_1 = 48$; $N_2 = 310$)	25.50	0.85	0.59	62.57	0.77
Stroop color ($N_1 = 48$; $N_2 = 309$)	25.50	0.79	0.64	66.11	0.75
Stroop words ($N_1 = 48$; $N_2 = 308$)	63.50	0.63	0.88	84.55	0.80
Stroop TT ($N_1 = 47$; $N_2 = 301$)	0.54	0.62	0.86	82.76	0.79
Stroop difference ($N_1 = 48$; $N_2 = 278$)	79.00	1.00	0.00	14.72	0.23
Planning Maze ($N_1 = 48$; $N_2 = 309$)	41.50	0.88	0.62	65.55	0.76
Planning Maze TT ($N_1 = 48$; $N_2 = 309$)	0.36	0.85	0.61	64.15	0.75
Interference C. I. ($N_1 = 48$; $N_2 = 309$)	25.50	0.73	0.71	71.15	0.77
Interference C.I. TT ($N_1 = 48$; $N_2 = 303$)	1.32	0.73	0.70	70.37	0.78

AKT Alters Konzentrations Test, *TT* total/time, *TMT B* Trail Making Test Version B, *C.I.* Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *SWT* Semantische Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *PWT* Phonematische Wortflüssigkeit, N_1 converters, N_2 non converters, *AUC* Area under curve.

The highest discrimination power with values over 0.80 for AUC was found in the domain memory. Here, all subtests reached very good to excellent diagnostic accuracy. These results of the memory domain agree with other studies investigating the predictive value of neuropsychological testing in detecting AD [7, 37, 42]. The results of the present study together with those of earlier studies support the assumption that verbal memory is affected very early. It is interesting to note that for the measures of TMT B – TMT A and the difference of Stroop color words, predictive power was very low. As both measures represent a difference score, the usefulness of difference scores is debatable. Future studies should look into quotient scores in terms of their usefulness.

Looking at the sensitivity of the NTBV, eight subtests (AKT, C.I.

Table 5b
Results of analyses of positive and negative predicted values and likelihood ratios at the chosen cut-off value with 95% confidence intervals.

Predictor variable	Cut off	PPV	NPV	LR+	LR-
Attention					
AKT ($N_1 = 48$; $N_2 = 310$)	36.50	0.25	0.96	2.13	0.27
AKT TT ($N_1 = 48$; $N_2 = 310$)	1.38	0.29	0.95	2.58	0.35
Digit-Symbol ($N_1 = 48$; $N_2 = 310$)	36.50	0.26	0.95	2.28	0.37
TMT B ($N_1 = 48$; $N_2 = 310$)	113.50	0.28	0.96	2.53	0.25
C.I. Symbols ($N_1 = 48$; $N_2 = 309$)	19.63	0.21	0.95	1.66	0.33
TMT B – TMT A ($N_1 = 47$; $N_2 = 310$)	240.50	0.13	100.00	1.01	0.00
Language					
SWT ($N_1 = 48$; $N_2 = 310$)	42.50	0.34	0.93	3.28	0.46
BNT ($N_1 = 48$; $N_2 = 309$)	13.50	0.26	0.90	2.28	0.68
Memory					
VSRT immediate ($N_1 = 48$, $N_2 = 310$)	5.50	0.40	0.93	4.36	0.46
VSRT total ($N_1 = 48$; $N_2 = 310$)	41.50	0.33	0.99	3.23	0.09
VSRT delayed ($N_1 = 48$; $N_2 = 309$)	7.50	0.39	0.98	4.07	0.13
VSRT recognition ($N_1 = 48$; $N_2 = 309$)	13.25	0.43	0.95	4.90	0.32
Executive function					
TMT A ($N_1 = 48$; $N_2 = 310$)	53.50	0.30	0.92	2.78	0.53
PWT ($N_1 = 48$; $N_2 = 309$)	21.50	0.25	0.90	2.18	0.72
5 Point ($N_1 = 48$; $N_2 = 310$)	25.50	0.24	0.96	2.08	0.25
Stroop color ($N_1 = 48$; $N_2 = 309$)	25.50	0.25	0.95	2.20	0.33
Stroop words ($N_1 = 48$; $N_2 = 308$)	63.50	0.45	0.94	5.20	0.43
Stroop TT ($N_1 = 47$; $N_2 = 301$)	0.54	0.41	0.94	4.42	0.45
Stroop difference ($N_1 = 48$; $N_2 = 278$)	79.00	0.15	100.00	1.00	0.00
Planning Maze ($N_1 = 48$; $N_2 = 309$)	41.50	0.26	0.97	2.31	0.20
Planning Maze TT ($N_1 = 48$; $N_2 = 309$)	0.36	0.25	0.96	2.18	0.24
Interference C. I. ($N_1 = 48$; $N_2 = 309$)	25.50	0.28	0.94	2.50	0.38
Interference C.I. TT ($N_1 = 48$; $N_2 = 303$)	1.32	0.28	0.94	2.43	0.39

AKT Alters Konzentrations Test, *TT* total/time, *TMT B* Trail Making Test Version B, *C.I.* Cerebral Insufficiency, *TMT A* Trail Making Test Version A, *SWT* Semantische Wortflüssigkeit, *BNT* Boston Naming Test, *VSRT* Verbal selective reminding test, *PWT* Phonematische Wortflüssigkeit, N_1 converters, N_2 non converters, *PPV* Positive predicted value, *NPV* Negative predicted value, *LR+* Positive likelihood ratio, *LR-* Negative likelihood ratio.

Symbols, TMT B, VSRT total recall, VSRT delayed recall, Five-Point Test, Planning Maze Test, Planning Maze total/time) reached a sensitivity over 0.80. A high sensitivity of a diagnostic test is important to avoid missing patients who need the treatment. The specificity of seven subtests (SWT, PWT, BNT, VSRT immediate recall, VSRT recognition, Stroop color words – words, Stroop total/time) approached values over 0.80. Thus, the probability of testing negative when the disease is absent is above 80.0% using these tests. The probability of not having AD-dementia if the test result is negative for the disease is high for all subtests ($0.90 \leq$ NPV), showing that it is easier to predict non-converters, while the probability of having AD if the NTBV subtests show a positive result can be interpreted as weak, indicating the difficulty of individual prediction (PPV < 0.50) [43]. Overall, recall and delayed-memory tasks can be seen as good predictors for conversion to dementia, indicating that it is easier to predict non-converters.

Poor values for discrimination accuracy for some subtests were determined. One reason may be the given time interval between both examinations. The progression of dementia evolves over several years. The large range of 12 – 48 months can introduce different progressions regarding the severity of the disease. Thus, stepwise logistic regression was used to survey the discrimination ability between converters to AD and non-converters of the NTBV subtest. Six subtests (VSRT delayed recall, VSRT total recall, VSRT recognition, TMT B, PWT, Stroop color words – words) were significant. This means that all six subtests provide enough information to allow a correct classification for each patient. Regarding the six subtests, 92.1% of patients in the total sample were diagnosed correctly, while only 7.9% received a false diagnosis. The results are consistent with the results of AUC, where all six subtests except for the PWT reached good to excellent diagnostic accuracy. Total internal consistency can be reported as strong with a Cronbach's α of 0.80 [45]. Therefore, from the aspect of internal consistency, it may be concluded that the NTBV is a reliable measurement.

NTBV scores in the total sample group were significantly lower for most of the subtests at the follow-up examination, representing disease progression. An excellent test-retest reliability can be assumed at $r \geq 0.80$ with a time interval of four weeks [46,47]. The test-retest reliability in the present study ranged between $r = 0.31$ and $r = 0.87$ across all diagnostic groups. This means that the test scores of the patients changed over time, indicating different degrees of change of cognition in different aspects of cognition. As already mentioned, one reason for the different test-retest reliabilities could be the long time between the first and second examinations. This may indicate that simple test-taking "practice" effects were minimized. Low test-retest reliability may thus indicate "real" changes in the cognitive function being assessed. Reliability will be higher when the interval between both examinations is rather short. A previous study came to this conclusion by examining the VSRT subtests [28].

While interpreting the results of the study, some limitations should be considered. The results cannot be extended to the general population due to the selective sample, which consisted solely of people who were referred from a doctor or actively made an appointment because of cognitive complaints. Moreover, there is no information about cognitive training or other treatments to maintain cognitive functions and delay the progression to AD between the intervals of both examinations [48–50]. This should be considered in future research. Despite the limitations, one of the strengths of the study is the rather large sample size, which covered a range from cognitively healthy to cognitively impaired. Furthermore, all patients received a thorough neuropsychological examination. Therefore, detailed data and accurate diagnoses were provided.

In conclusion, the study demonstrates that the NTBV may be successfully used as a predictive measure to assess cognitive functions in patients with different neurocognitive status. Future research should focus on subtests, as they have a good predictive value for diagnosing AD-dementia. This may increase predictive power and guarantee overall test efficiency.

Data availability statement

Data available on request from the authors.

Transparency statement

Assoc. Prof. Priv. Doz. Mag. Dr. Johann Lehrner affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Ethical statement

The dataset was collected from 2008 to May 2018 within a current research project: the Vienna Conversion to Dementia Study. The study was approved by the Ethical Committee of the Medical University of Vienna.

CRediT authorship contribution statement

Anna Garcia Rosas: Conceptualization, Formal analysis, Investigation, Writing – original draft. **Elisabeth Stögmann:** Investigation, Writing – review & editing. **Johann Lehrner:** Conceptualization, Investigation, Project administration, Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgements

The NTBV can be obtained from www.psimistri.com – global psychometric test and intervention systems.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dscb.2021.100028.

References

[2] C.L. Satizabal, A.S. Beiser, V. Chouraki, G. Chêne, C. Dufouil, S. Seshadri, Incidence of dementia over three decades in the Framingham heart study, N. Engl. J. Clin. Med. 374 (6) (2016) 523–532, https://doi.org/10.1056/NEJMoa1504327.

[4] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-4), American Psychiatric Association, Washington, DC, 2013.

[5] F. Jessen, B. Wiese, C. Bachmann, et al., Prediction of dementia by subjective memory impairment effects of severity and temporal association with cognitive impairment, Arch. Gen. Psychiatry 67 (4) (2010) 414–422, https://doi.org/10.1001/archgenpsychiatry.2010.30.

[6] F. Jessen, S. Wolfsgruber, B. Wiese, et al., AD dementia risk in late MCI, in early MCI, and in subjective memory impairment, Alzheimers Dement 10 (1) (2014) 76–83, https://doi.org/10.1016/j.jalz.2012.09.017.

[7] J. Lehrner, R. Gufler, G. Guttmann, J. Maly, A. Gleiß, E. Auff, P. Dal-Bianco, Annual conversion to Alzheimer 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, Wien Klin Wochenschr 117 (18) (2005) 629–635, https://doi.org/10.1007/s00508-005-0428-6.

[8] C. Drummond, G. Coutinho, R.P. Fonseca, et al., Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment, Front. Aging Neurosci. 7 (2015) 1–11.

[9] M.F. Elias, A. Beiser, P.A. Wolf, R. Au, R.F. White, R.B. D'Agostino, The preclinical phase of Alzheimer disease: a 22-year prospective study of the Framingham cohort, Arch. Neurol. 57 (6) (2000) 808–813, https://doi.org/10.1001/archneur.57.6.808.

[10] E. Grober, R.B. Lipton, C. Hall, H. Crystal, Memory impairment on free and cued selective reminding predicts dementia, Neurology 54 (4) (2000) 827–832, https://doi.org/10.1212/WNL.54.4.827.

[11] A. Maseda, L. Lodeiro-Fernández, L. Lorenzo-López, L. Núñez-Naveira, A. Balo, J. C. Millán-Calenti, Verbal fluency, naming and verbal comprehension: three aspects of language as predictors of cognitive impairment, Aging Ment. Health 18 (8) (2014) 1037–1045.

[12] C. Schiffczyk, B. Romero, C. Jonas, C. Lahmeyer, F. Müller, M.W. Riepe, Efficacy of short-term inpatient rehabilitation for dementia patients and caregivers: prospective cohort study, Dement Geriatr. Cogn. Disord. 35 (5–6) (2013) 300–312.

[13] G. Pusswald, D. Moser, A. Gleiß, S. Janzek-Hawlat, E. Auff, P. Dal-Bianco, J. Lehrner, Prevalence of mild cognitive impairment subtypes in patients attending a memory outpatient clinic-comparison of two modes of mild cognitive impairment classification. Results of the Vienna conversion to dementia study, Alzheimers Dement 9 (4) (2013) 366–376, https://doi.org/10.1016/j.jalz.2011.12.009.

[14] J. Lehrner, J. Maly, A. Gleiß, E. Auff, P. Dal-Bianco, The Vienna neuropsychological test battery (VNTB) for detecting dementia: standardization, norms, and validation, Psychol. Österreich 4 (2007) 358–365.

[15] F. Jessen, R.E. Amariglio, M. Van Boxtel, et al., A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease, Alzheimers Dement 10 (6) (2014) 844–852, https://doi.org/10.1016/j.jalz.2014.01.001.

[16] R.C. Petersen, Mild cognitive impairment as a diagnostic entity, J. Intern. Med. 256 (3) (2004) 183–194, https://doi.org/10.1111/j.1365-2796.2004.01388.x.

[17] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, E.M. Stadlan, 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 (7) (1984) 939, https://doi.org/10.1212/WNL.34.7.939.

[18] Schmidt K.H., Metzler P. Wortschatztest: WST. Weinheim: Beltz; 1992.

[19] M. Hautzinger, F. Keller, C. Kühner, Beck Depressions-Inventar (BDI-II), Harcourt Test Services, Frankfurt, 2006.

[20] M.F. Folstein, S.E. Folstein, P.R. Mc Hugh, Mini-mental state ": a practical method for grading the cognitive state of patients for the clinician, J. Psychiatr. Res. 12 (3) (1975) 189–198.

[21] G. Gatterer, Alters-Konzentrations-Test (AKT), Hogrefe, Göttingen, 1990.

[22] U. Tewes, Hamburg-Wechsler-Intelligenztest Für Erwachsene – Revision 1991 (HAWIE-R), Hans Huber, Bern, 1994.

[23] S. Lehrl, B. Fischer, Kurztest Für Cerebale Insuffizienz (c.I.-Test), Vless, Ebersberg, 1997.

[24] R. Reitan, Trail Making Test (TMT), Reitan Neuropsychology Laboratory, Tucson, 1979.

[25] M. Regard, E. Strauss, P. Knapp, Children's production on verbal and non-verbal fluency tasks, Percept. Mot. Skills 55 (3) (1982) 839–844, https://doi.org/10.2466/pms.1982.55.3.839.

[26] W.D. Oswald, U.M. Fleischmann, Das Nünberger-Alters-Inventar, Hogrefe Verlag, Göttingen, 1997.

[27] Goodglass H., Kaplan P. The assessment of aphasia and related disorders. Philadelphia: Lea & Fabinger; 1983.

[28] J.C. Morris, A. Heyman, R.C. Mohs, et al., The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease, Neurology 39 (9) (1989) 1159–1165.

[29] J. Lehrner, A. Gleiß, J. Maly, E. Auff, P. Dal-Bianco, Der verbale selektive reminding test (VSRT). Ein Verfahren zur überprüfung verbaler gedächtnisfunktionen, Neuropsychiatr 20 (3) (2006) 204–214.

[30] T. Crook, M. Salama, J. Gobert, A computerized test battery for detecting and assessing memory disorders, in: A Bes, J Cohn, S Hoyer, JP Marc-Vergenes, HM Wisniewski (Eds.), Senile Dementias: Early Detection, John Libbey Eurotext, London-Paris, 1986, pp. 79–85.

[31] J.R. Youngjohn, G.J. Larrabee, T.H. Crook, First-last names and the grocery list selective reminding test: two computerized measures of everyday verbal learning, Arch. Clin. Neuropsychol. 6 (4) (1991) 287–300, https://doi.org/10.1016/0887-6177(91)90006-U.

[32] D. Hitzl, Neuropsychologischer Outcome Nach Tiefer Hirnstimulation bei Parkinsonpatienten – Eine Retrospektive Datenauswertung, Medical University of Vienna, 2013. Unpublished manuscript.

[33] Macher S. Reliabilität und validität der neuropsychologischen testbatterie Vienna bei patienten mit leichter demenz – eine pilotstudie. Medical University of Vienna; 2013. Unpublished manuscript.

[34] K.D. Kubinger, Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens, Hogrefe Verlag, Vienna, 2009.

[35] Moser S. Cognitive functions and mortality in patients with subjective cognitive decline, mild cognitive impairment and Alzheimer's disease. Medical University of Vienna; 2018. Unpublished manuscript.

[36] J. Cohen, A power primer, Psychol. Bull. 122 (1) (1992) 155–159.

[37] J. Lehrner, B. Bodendorfer, C. Lamm, D. Moser, P. Dal-Bianco, E. Auff, G Pusswald, Subjective memory complaints and conversion to dementia in patients with subjective cognitive decline and patients with mild cognitive impairment, Z Neuropsychol 27 (2) (2016) 85–93.

[38] N. Valencia, J. Lehrner, Screening for dementia with the Vienna visuo-constructional test 3.0 screening (VVT 3.0 screening), Neuropsychiatrie 32 (4) (2018) 196–203.

[39] A.J. Mitchell, H. Beaumont, D. Ferguson, M. Yadegarfar, B. Stubbs, Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis, Acta Psychiatr. Scand. 130 (6) (2014) 439–451, https://doi.org/10.1111/acps.12336.

[40] G.M. McKhann, D.S. Knopman, H. Chertkow, et al., The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, Alzheimers Dement 7 (3) (2011) 263–269.

[41] R.A. Sperling, P.S. Aisen, L.A. Beckett, et al., Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, Alzheimers Dement 7 (3) (2011) 280–292.

[42] P. Chen, G. Ratcliff, S.H. Belle, J.A. Cauley, S.T. DeKosky, M. Ganguli, Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented, Neurology 55 (12) (2000) 1847–1853.

[43] R. Jaeschke, G.H. Guyatt, D.L. Sackett, et al., Users' guides to the medical literature: III. How to use an article about a diagnostic test B. What are the results and will they help me in caring for my patients? JAMA 271 (9) (1994) 703–707.

[45] D. George, M. Mallery, Using SPSS For Windows step By Step: A Simple Guide and Reference 11.0 Update, Allyn & Bacon, Boston, 2003.

[46] C.W. Chen, H. Chu, C.F. Tsai, et al., The reliability, validity, sensitivity, specificity and predictive values of the Chinese version of the Rowland universal dementia assessment scale, J. Clin. Nurs. 24 (21–22) (2015) 3118–3128.

[47] M.M. Gonçalves, M.S. Pinho, M.R. Simões, Test–retest reliability analysis of the Cambridge neuropsychological automated tests for the assessment of dementia in older people living in retirement homes, Appl. Neuropsych.-Adul. 23 (4) (2016) 251–263.

[48] I. Carrière, A. Fourrier-Reglat, J.F. Dartigues, O. Rouaud, F. Pasquier, K. Ritchie, M. L. Ancelin, Drugs with anticholinergic properties, cognitive decline, and dementia in an elderly general population: the 3-city study, Arch. Intern. Med. 169 (14) (2009) 1317–1324.

[49] N.T. Hill, L. Mowszowski, S.L. Naismith, V.L. Chadwick, M. Valenzuela, A. Lampit, Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis, Am. J. Psychiatry 174 (4) (2016) 329–340.

[50] T. Ngandu, J. Lehtisalo, A. Solomon, et al., A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial, Lancet Glob. Health 385 (9984) (2015) 2255–2263.