

This is an Open Access article licensed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Unported license (CC BY-NC) (www.karger.com/OA-license), applicable to the online version of the article only. Distribution permitted for non-commercial purposes only.

Original Research Article

The Extension of the German CERAD Neuropsychological Assessment Battery with Tests Assessing Subcortical, Executive and Frontal Functions Improves Accuracy in Dementia Diagnosis

Nicole S. Schmid^a Michael M. Ehrensperger^a Manfred Berres^b
Irene R. Beck^a Andreas U. Monsch^a

^aMemory Clinic Basel, University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland; ^bDepartment of Mathematics and Technology, RheinAhrCampus, Remagen, Germany

Key Words

Alzheimer's disease · Diagnostic equipment · Frontotemporal dementia · Vascular dementia · Neuropsychological tests · Aging

Abstract

Background/Aims: Alzheimer's disease (AD) is the most common form of dementia. Neuropsychological assessment of individuals with AD primarily focuses on tests of cortical functioning. However, in clinical practice, the underlying pathologies of dementia are unknown, and a focus on cortical functioning may neglect other domains of cognition, including subcortical and executive functioning. The current study aimed to improve the diagnostic discrimination ability of the Consortium to Establish a Registry for Alzheimer's Disease – Neuropsychological Assessment Battery (CERAD-NAB) by adding three tests of executive functioning and mental speed (Trail Making Tests A and B, S-Words). **Methods:** Logistic regression analyses of 594 normal controls (NC), 326 patients with mild AD and 224 patients with other types of dementia (OD) were carried out, and the area under the curve values were compared to those of CERAD-NAB alone. **Results:** All comparisons except AD-OD (65.5%) showed excellent classification rates (NC-AD: 92.7%; NC-OD: 89.0%; NC-all patients: 91.0%) and a superior diagnostic accuracy of the extended version. **Conclusion:** Our findings suggest that these three tests provide a sensible addition to the CERAD-NAB and can improve neuropsychological diagnosis of dementia.

© 2014 S. Karger AG, Basel

Prof. Andreas U. Monsch
Memory Clinic Basel, University Center for Medicine of Aging
Felix Platter Hospital
Schanzenstrasse 55, CH–4031 Basel (Switzerland)
E-Mail andreas.monsch@unibas.ch

Introduction

Neuropsychological assessment plays a crucial role in differential dementia diagnosis [1]. Alzheimer's disease (AD) is the most common cause of dementia and primarily affects 'cortical' functioning such as verbal episodic memory. Thus, most neuropsychological test batteries focus on these cortical cognitive functions [2]. However, in clinical practice, many types of distinct deficits may be observed and multiple hypotheses must be evaluated. Solely focusing on cortical functioning may fail to identify other impairments that may be indicative of dementia from etiologies other than AD. The goal of the current study was to determine if the inclusion of tests of subcortical and frontal functioning could improve the diagnostic accuracy of the Consortium to Establish a Registry for Alzheimer's Disease – Neuropsychological Assessment Battery (CERAD-NAB) [3], a commonly used neuropsychological test battery. Accurate clinical diagnosis of dementia is crucial for determining prognosis and implementing appropriate disease-specific treatments, also in due consideration of the enormous healthcare costs caused by dementia [4]. For diagnostic evaluation, recent EFNS-ENS Guidelines recommend that detailed cognitive and neuropsychological assessment is performed in all patients [5]. Much attention has been dedicated toward the development of instruments focusing on the detection of the earliest changes in cognitive functions affected in dementia [6, 7]. AD accounts for approximately 70% of all dementia cases and has been identified as a research priority [8]. This focus has led to an increased knowledge about the disease and to the identification of deficits in a broad range of cognitive domains. Impairment in learning, memory, attention, executive functioning, and language are some of the earliest observable cognitive signs of AD (for a review see [9]). Although AD is the leading cause of dementia in the elderly, dementia can arise from a variety of etiologically and neuropathologically distinct disorders, causing different patterns of cognitive impairment [2]. According to the actual EFNS-ENS Guidelines, the behavioral variant of frontotemporal dementia (FTD) is characterized by a progressive decline in interpersonal skills and emotional disorders such as apathy, emergence of a variety of abnormal behaviors and executive dysfunctions, while in subcortical forms of dementia, such as vascular dementia (VaD), disruption of subcortical frontal loops may lead to psychomotor slowing and executive dysfunctions [5]. In practice, clinicians must evaluate multiple possible etiologies when determining the potential presence and cause of neuropsychological dysfunction.

For example, different profiles of neuropsychological impairment can distinguish between AD and VaD, another common type of dementia. VaD refers to a cumulative decline in cognitive functioning secondary to cerebral infarction, ischemic injury, or hemorrhagic lesions [2]. The effects of cerebrovascular disease on cognition have recently been defined by the term vascular cognitive impairment [10]. The most prominent cognitive deficits observed in vascular cognitive impairment are executive/subcortical functions including goal formulation, initiation, planning, organization, shifting and psychomotor speed [11], sequencing, speed of mental processing, performance on unstructured tasks and attention [12], semantic memory, attentional abilities, visuospatial reasoning, and perceptual skills [13]. In the clinical setting, it is challenging to distinguish between VaD and AD because vascular risk factors and stroke may also be observed in individuals with AD. Vascular factors, such as cerebrovascular disease/strokes, and white matter lesions may trigger and/or modify the progression of AD [10]. Several studies [12, 14] compared the neuropsychological profiles in patients with subcortical VaD and patients with AD. The results revealed that memory impairment was more prominent than executive dysfunction in almost two thirds of AD patients, whereas patients with subcortical VaD generally showed greater deficits in executive functions than in verbal episodic memory. Thus, inclusion of tests examining subcortical/executive functions seems to be a promising strategy for distinguishing between impairment due to AD and VaD.

Similarly, tests targeting frontal functions may be helpful for neuropsychological test batteries aiming to distinguish between AD and FTD. FTD, for example, may make up 50% of dementia cases in adults younger than 60 years of age [15] and 6–12% of late-onset dementia cases [16]. Drawing a clinical distinction between FTD and AD can be difficult, particularly at the early stages of disease [17]. Previous studies [18, 19] have observed that patients with FTD are typically more impaired than patients with AD on spontaneous word generation tasks (i.e., letter and category fluency). In contrast, patients with FTD demonstrate lesser deficits on tests of memory and visuospatial abilities compared to patients with AD. Rascovsky et al. [20] were able to distinguish patients with FTD and AD using two verbal fluency measures. They observed that patients with FTD exhibit worse performance on letter fluency compared to semantic category fluency, whereas patients with AD perform worse on semantic category fluency compared to letter fluency. This contrasting pattern of fluency deficits in FTD and AD may indicate the differences in the relative contribution of frontal lobe-mediated retrieval deficits (more prominent in FTD) and temporal lobe-mediated semantic deficits (more prominent in AD) in the two disorders [2]. Inclusion of both semantic category fluency and letter fluency in a neuropsychological test battery might provide additional information about the underlying pathology and enhance the validity of dementia diagnosis.

In this study, we extend the well-established CERAD-NAB [3], German version, with tests assessing processing speed (Trail Making Test A, TMTA [21]), executive functioning (Trail Making Test B, TMTB [21]), and frontal functions (letter fluency: S-Words) [22] to determine whether these additions may enhance dementia diagnosis compared to the CERAD-NAB alone. The time needed for the administration of these additional tests will increase the neuropsychological assessment by approximately 10 min. We compared patients with AD to normal controls (NC) and also to other types of dementia, including VaD, FTD, and mixed dementia (MD). Our hypothesis is that the extended battery will result in superior discrimination between AD and NC and also between AD and other types of dementia. We will also illustrate the clinical utility of our findings by applying them to two cases and determine their probabilities of having dementia.

Materials and Methods

Participants

NC Group

The sample of cognitively healthy elderly individuals was selected from the BASEL (Basel Study on the ELderly) project [23], a longitudinal study that aims to identify preclinical markers of AD. This project was approved by our local ethics committee, and written informed consent was obtained from all participants. Eleven hundred participants qualified for inclusion in the current sample. The following inclusion and exclusion criteria were applied for study inclusion: (1) German as first language; (2) no more than one age-, gender-, and education-corrected Z-score ≤ -1.96 on any variable on the German version of the CERAD-NAB [24]; (3) absence of systemic or neurological illnesses according to self-reported medical history; (4) participants did not meet diagnostic criteria for a current major depressive episode according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) [25]; (5) participants did not demonstrate any visual or hearing impairment that could have adversely influenced the neuropsychological assessment; (6) a minimum age of 50 years; (7) at least 7 years of formal education, and (8) a Mini-Mental State Examination (MMSE) score [26] of ≥ 24 . Out of these 1,100 individuals, 594 participants completed TMTA, TMTB and S-Words and composed the NC sample for the current study. Demographic characteristics are displayed in table 1.

Table 1. Characteristics of the study sample

Groups	NC	AD	OD	OD subgroups		
				MD	VaD	FTD
Participants	594	326	224	116	69	39
Sex, male/female	331/263	137/189	123/101	63/53	36/33	24/15
Age, years	68.37±7.66	75.66±5.91	75.17±7.46	77.34±5.71	75.07±6.97	68.87±9.29
Education, years	12.32±3.00	11.83±3.05	12.31±3.18	12.44±3.30	12.06±3.06	12.36±3.10
MMSE score [24]	28.9±1.14	26.2±1.60	26.5±1.60	26.3±1.58	26.7±1.53	26.6±1.58

Values are numbers or mean ± SD.

AD Patients with Very Mild to Mild Dementia

Patients were chosen from the Memory Clinic, University Hospital, Basel, Switzerland, where they had been diagnosed with probable AD according to the criteria of the National Institute for Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [27] and of the DSM-IV [25]. The multidisciplinary assessment included (1) independent, detailed clinical interviews with the patient and (where possible) informant formally assessing patients' mood and cognitive and everyday functioning; (2) thorough medical examinations including neurological status; (3) an MRI examination with sequences optimized to detect structural brain abnormalities associated with cognitive impairment; (4) a complete laboratory workup, and (5) a comprehensive neuropsychological testing including the German version of the CERAD-NAB plus TMTA and S-Words. Diagnoses were established at an interdisciplinary diagnosis conference with geriatricians, neurologists, neuropsychologists, a neuropathologist, neuroimaging experts, and psychiatrists (for details of these assessments and procedures, see [28]). Since CERAD-NAB data were available at the diagnosis conference, the question of circularity arises, i.e. the concern that the reported correct classification rate (CCR) of the CERAD-NAB was artificially inflated because patients may have been a priori defined as individuals with lower CERAD-NAB scores. However, diagnoses were never based on the neuropsychological assessment alone, but rather on the integration of all comprehensive multidisciplinary findings, which cannot logically be equated with a particular CERAD-NAB score. Nevertheless, we cannot completely rule out that the reported CCR was artificially inflated by the availability of CERAD-NAB scores at the diagnosis conference. The group comprised 326 patients with AD. Demographic characteristics are shown in table 1. All participants scored at 24 or higher on the MMSE, ensuring that participants were in a very early stage of the disease [24].

Patients with Other Types of Dementia

In order to gain a sample with types of dementia other than AD (OD), we included 224 individuals who obtained a diagnosis of a dementia with a subcortical component at the Memory Clinic Basel. Diagnosis was established as noted above. Characteristics of the study sample are displayed in table 1. Three subgroups were defined within the OD group.

One hundred and sixteen individuals received a diagnosis of MD. These patients received a diagnosis of probable AD according to NINCDS-ADRDA and DSM-IV criteria and also met criteria for a subcortical component of dementia according to Erkinjuntti et al. [11]. The diagnosis of MD necessarily includes a cognitive syndrome (both memory impairment and dysexecutive syndrome) with evidence of a relevant cerebrovascular disease by brain imaging and presence or history of neurologic signs as evidence for cerebrovascular disease.

Ninety-six patients were diagnosed with probable VaD according to criteria set forth by Roman et al. [29]. These criteria included (1) dementia (decline in memory and deficits in at least one additional domain), and (2) existence of a cerebrovascular disease (focal neurological signs consistent with stroke) with supportive evidence from a computed tomography or MRI scan.

Thirty-nine patients were diagnosed with FTD using criteria set forth by Neary et al. [30]. Core diagnostic features essentially included (1) insidious onset and gradual progression; (2) early decline in social and interpersonal conduct in the course of the disease; (3) early impairment in regulation of personal conduct; (4) early emotional blunting, and (5) early loss of insight.

Cognitive Tests

We used Z-scores adjusted for age, gender, and education [31, 32] for all measures.

CERAD-NAB

The German version of the CERAD-NAB consists of nine different tasks measuring cognitive performance in domains including memory, language, praxis, and orientation. The current study added three additional tests examining frontal, subcortical, and executive functioning to the CERAD-NAB. We will henceforth refer to this expanded test battery as CERAD-NAB Plus, consisting of 12 variables.

Frontal executive functions were measured using S-Words, a task measuring the ability of participants to generate as many words as possible beginning with the given letter 'S' within 60 s. In this measure, the following restrictions are given to participants: any words are allowed except for proper nouns, numbers, composite words, or words that are the same as a previously used word with a different tense or suffix. Performance is evaluated with the total number of words listed.

Subcortical functioning was evaluated using TMTA, a measure of psychomotor speed and visual scanning. In this measure, participants must draw lines connecting the numbers 1 through 25 in ascending order as quickly as possible. The numbers are distributed over a sheet of paper. Performance is evaluated using the total completion time. A maximum of 180 s is allowed. Errors are not included in evaluating scores; however, all errors are corrected, which may increase the completion time for the task.

Executive functioning was evaluated using TMTB, a measure of mental flexibility and the ability to switch between two different modalities. In TMTB, both numbers (1–13) and letters (A–L) are distributed over a sheet of paper. Participants are asked to draw lines to connect numbers and letters in ascending pattern, but with the added task of alternating between modalities (i.e., 1-A-2-B-3-C, etc.). Again, performance is evaluated using the total completion time in seconds. A maximum of 300 s is allowed.

Missing Values in the TMT

The sample contained missing values in the TMT. The reasons for missing values included premature interruption of the test by the participant or examiner, inadequate comprehension of test instructions, and excess of the maximal time allowed. Detailed information about distribution of missing data can be found in table 2.

TMTB data were missing for 51.9% of the AD sample and 44.6% of the OD patients. Considering this large amount of missing data, we searched for an appropriate method to replace the values. Since the values were missing due to censoring (see above), neither replacement by a mean value nor replacement by the lowest value of 300 s is acceptable; both methods would result in a strong bias. Instead, we estimated a penalty that can be applied to the individual Z-scores of the noncompleters at the lowest value of 300 s. In order to create

Table 2. Number and percentages of missing values in the TMTs

Groups	Missing values in TMTA [19]	Missing values in TMTB [19]
NC (n = 594)	0	7 (1.1%)
AD (n = 326)	13 (4.0%)	169 (51.8%)
OD (n = 224)	6 (1.8%)	100 (44.6%)

such a penalty, we focused on those subjects who completed the TMTB in the time interval between 240 and 300 s (needing at least 80% of the maximum time allowed) and those subjects who did not complete the TMTB. We aimed to compare their respective sample means to a representative score for test performance. We chose the z-values of the CERAD total score [24]. The slowest 20% of the sample of completers showed a mean of -2.41 , while the noncompleters achieved a mean of -3.24 . The resulting difference of 0.83 is a possible penalty that could be derived from the Z-score equal to the worst performance in the TMTB in each participant who failed in this test. The same procedure has been applied for the interval between 270 and 300 s (the slowest 10% of the sample of completers, mean = -2.54), with a possible penalty of 0.6 . We decided to subtract a ‘penalty’ of 0.7 from the patient’s Z-score for TMTB of 300 s. This procedure ensures an individual, demographically adjusted penalty for individuals who could not complete the task. The same procedure was carried out for missing values on TMTA, resulting in a penalty of 0.4 .

Statistical Analysis

The first aim of the study was to get information about the ability of CERAD-NAB alone and CERAD-NAB Plus to correctly differentiate between patients with different kinds of dementia and NC. A comparison of their discrimination abilities was performed. The following groups were compared using respective cutoff values (share of patients in the entire sample): (1) NC versus AD to gain information about the difference of correct classification in NC and cortical dementia; (2) NC versus OD to examine classification between NC and frontal/subcortical dementia and NC; (3) NC versus all dementias (DEM) to provide information about correct classification of NC and all types of dementia, and (4) AD versus OD to show the contribution of the ‘plus’ tests (TMTA, TMTB, and S-Words) in differentiating between cortical and subcortical/frontal dementia. Logistic regression analyses were performed on all 9 or 12 CERAD-NAB or CERAD-NAB Plus scores, respectively. This method assigns an optimal weight to each variable in the model in order to achieve the best CCRs and corresponding sensitivities and specificities. However, due to intercorrelations in the predictors, some regression coefficients were assigned coefficients with an algebraic sign inconsistent with the direction predicted by theory (i.e., a higher Z-score predicted a higher probability of dementia). Therefore, a penalization method was applied. In clinical practice usually all tests from the CERAD-NAB are carried out. In order to gain information about the influence of every single variable, we chose the penalization by ridge regression. By doing so, all regression coefficients are reduced but remain in the model. The optimal penalization parameter lambda was estimated by cross-validation, using the R package glmnet. The obtained coefficients are more stable and were applied to the dataset, resulting in a prediction score for each group comparison. Based on the receiver operating characteristic (ROC), optimal CCR and the corresponding sensitivity and specificity were determined for each model. The discrimination abilities of the two test batteries were compared using the area under the curve (AUC) method according to DeLong et al. [33], as implemented in the R package pROC [34].

A concern of this study was to illustrate the implementation of our findings in clinical practice by applying them to randomly selected patients and computing their probability of

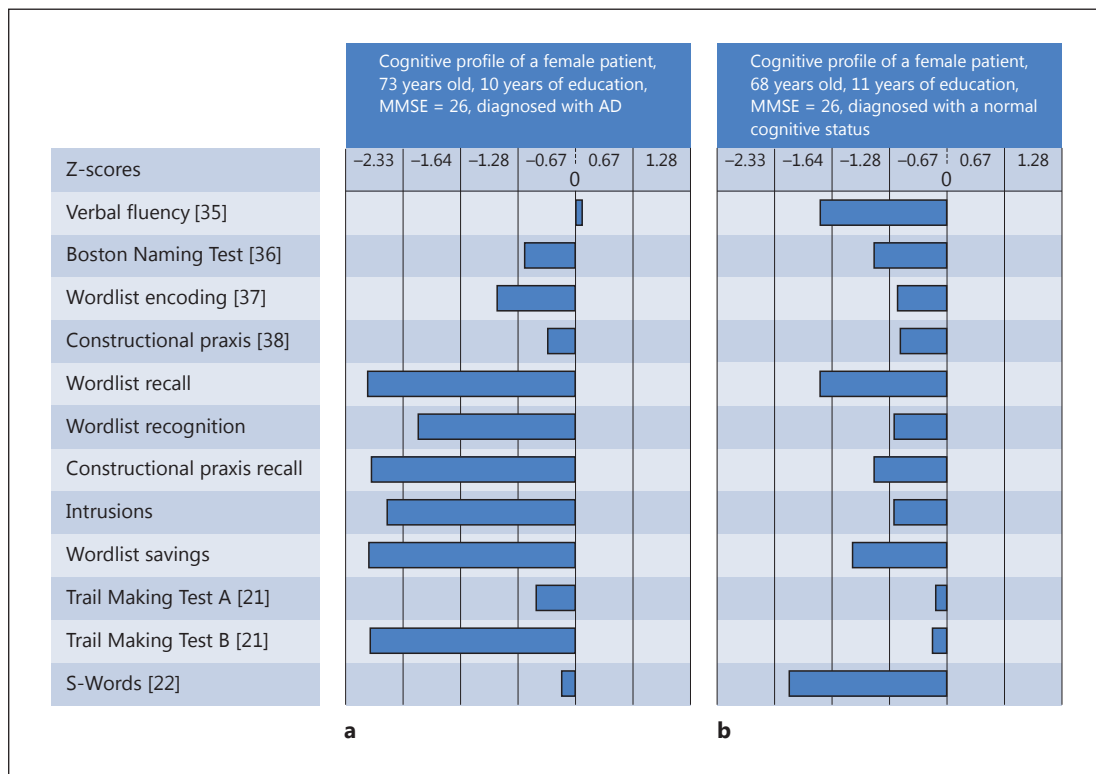


Fig. 1. CERAD-NAB Plus data from patients examined in the Memory Clinic Basel.

having dementia. Regression weights from the ridge regression models distinguishing between NC and demented patients were applied to a randomly chosen patient from the Memory Clinic Basel diagnosed with Alzheimer’s dementia (fig. 1a) and to a demographically comparable healthy individual outside the study sample (fig. 1b). Then, we estimated their probability of being demented by applying the logistic function to the prediction. A high score corresponds to a high probability of being demented. The probability cutoff for distinguishing between the two groups was chosen as the proportion of the sample size.

Results

Optimal weights for each variable gained from ridge regression to distinguish best between groups are displayed in table 3. The applied coefficients from ridge regression (table 3) resulted in a prediction score for each group comparison. By producing ROC of this score, the optimal CCR and corresponding sensitivity and specificity were conducted. The AUC is presented as well for each model. Results are shown in table 4.

Both CERAD-NAB and CERAD-NAB Plus show an excellent ability to distinguish correctly between NC and patients suffering from AD. A correct discrimination of >90% could be obtained when diagnosing NC and all demented patients with both batteries. CCR of NC and other types of dementia is in general slightly lower than in the antecedent comparisons for both batteries but still satisfying (87.55 vs. 88.95%). Comparison of AD versus OD revealed no relevant contribution to a satisfactory classification (48.2%) of groups in both CERAD-NAB and CERAD-NAB Plus.

Table 3. Coefficients obtained from ridge regression for each variable of CERAD-NAB and CERAD-NAB Plus that differentiate best between the observed groups

Groups	CERAD-NAB	CERAD-NAB Plus
NC-AD (cutoff: 0.35)	$\lambda = 0.56$	$\lambda = 0.07$
Verbal fluency (animal naming)	-0.46	-0.34
Boston Naming Test (15 items)	-0.22	-0.15
Wordlist encoding	-0.48	-0.37
Constructional praxis	-0.06	0.03
Wordlist recall	-0.45	-0.37
Wordlist recognition	-0.27	-0.24
Constructional praxis recall	-0.35	-0.28
Intrusions	-0.21	-0.16
Wordlist savings	-0.14	-0.14
TMTA		-0.16
TMTB		-0.38
S-Words		-0.06
NC-OD (cutoff: 0.27)	$\lambda = 0.10$	$\lambda = 0.07$
Verbal fluency (animal naming)	-0.49	-0.44
Boston Naming Test (15 items)	-0.18	-0.14
Wordlist encoding	-0.37	-0.33
Constructional praxis	0.09	0.04
Wordlist recall	-0.33	-0.33
Wordlist recognition	-0.08	-0.05
Constructional praxis recall	-0.31	-0.26
Intrusions	-0.16	-0.14
Wordlist savings	-0.08	-0.07
TMTA		-0.21
TMTB		-0.41
S-Words		-0.07
NC-DEM (cutoff: 0.48)	$\lambda = 0.005$	$\lambda = 0.06$
Verbal fluency (animal naming)	-0.56	-0.43
Boston Naming Test (15 items)	-0.20	-0.14
Wordlist encoding	-0.44	-0.36
Constructional praxis	-0.08	0.03
Wordlist recall	-0.41	-0.36
Wordlist recognition	-0.18	-0.16
Constructional praxis recall	-0.37	-0.30
Intrusions	-0.22	-0.17
Wordlist savings	-0.10	-0.10
TMTA		-0.18
TMTB		-0.43
S-Words		-0.05
AD-OD (cutoff: 0.59)	$\lambda = 0.44$	$\lambda = 0.40$
Verbal fluency (animal naming)	0.10	0.10
Boston Naming Test (15 items)	-0.01	-0.01
Wordlist encoding	-0.10	-0.11
Constructional praxis	0.02	0.03
Wordlist recall	-0.11	-0.11
Wordlist recognition	-0.12	-0.12
Constructional praxis recall	-0.06	-0.07
Intrusions	-0.04	-0.04
Wordlist savings	-0.04	-0.04
TMTA		0.04
TMTB		-0.01
S-Words		0.07

Cutoff values (share of patients in the entire sample) are provided.

Table 4. Sensitivity, specificity, CCR and AUC of the CERAD-NAB and CERAD-NAB Plus for observed group comparisons

Statistical values	CERAD-NAB, %	CERAD-NAB Plus, %
NC-AD (cutoff: 0.35)		
Sensitivity	95.1	92.6
Specificity	90.1	92.8
CCR	92.6	92.7
AUC	0.974	0.978
NC-OD (cutoff: 0.27)		
Sensitivity	87.9	92.0
Specificity	87.2	85.9
CCR	87.6	89.0
AUC	0.941	0.953
NC-DEM (cutoff: 0.48)		
Sensitivity	90.4	88.9
Specificity	90.2	93.1
CCR	90.3	91.0
AUC	0.960	0.967
AD-OD (cutoff: 0.59)		
Sensitivity	84.0	83.1
Specificity	48.2	47.8
CCR	66.1	65.6
AUC	0.706	0.712

We compared the AUC values between the two batteries using procedures set forth by DeLong et al. [33] to determine whether the CERAD-NAB Plus was significantly better at discriminating between the groups than the CERAD-NAB alone. The results revealed that CERAD-NAB Plus discriminates significantly better than CERAD-NAB alone between NC and AD ($Z = -2.34$, $p = 0.0194$), NC and OD ($Z = -3.12$, $p = 0.0018$), and NC and DEM ($Z = -3.55$, $p = 0.0004$). CERAD-NAB Plus did not distinguish AD from OD significantly better than CERAD-NAB alone ($Z = -1.18$, $p = 0.2398$).

In order to illustrate the practical application of the extended test battery, we applied the weights obtained from the group comparison between NC and demented patients to the CERAD-NAB Plus scores of a randomly selected patient diagnosed with AD in the Memory Clinic Basel (fig. 1a). First, we replaced the missing value of TMTB by her Z-score for TMTB of 300 s. We deducted a penalty of 0.7, resulting in a Z-score of -2.84 . Applied to the model, a prediction score of 2.57 resulted and revealed a probability of 92.89% of being demented. The same procedure has been applied to a woman diagnosed as cognitively healthy in the Memory Clinic Basel (fig. 1b) with similar characteristics; the outcome was a prediction score of -1.10 and a probability of 24.97% of being in the demented group.

Discussion

In addition to domains tested in the well-established CERAD-NAB, the newly created CERAD-NAB Plus especially assesses psychomotor speed, executive and frontal functions known to be affected particularly in subcortical/frontal forms of dementia. The present study provides evidence for an improved correct classification of patients suffering from any type of dementia and NC in the CERAD-NAB Plus compared to CERAD-NAB alone. No significant differentiation between cortical and subcortical/frontal forms of dementia could be shown in both CERAD-NAB and CERAD-NAB Plus.

As a main finding, the current study emphasizes the addition of selected tests to the CERAD-NAB in order to improve dementia diagnosis. These tests do not enhance the differential diagnosis of dementia. Verbal fluency (animal naming) and TMTB appeared to be the most valuable variables in all comparisons except AD versus OD. Impaired verbal fluency was already described in many influential longitudinal studies searching for preclinical markers of dementia 4 (Bronx Aging Study [39]), 5 (Canadian Study of Health and Aging [40]), 9 (Personnes Agées QUID [41]) and even 12 years (PAQUID [42]) before the onset of observable clinical symptoms. Impairment on TMTB appeared likewise amongst the earliest preclinical signs either in combination with impaired verbal fluency such as in the Baltimore Longitudinal Study of Aging [43] and in the Cardiovascular Health Study [44] or isolated in the Monogahela Valley Independent Elders Survey [45]. Two meta-analyses [40, 46] have stated that impairments in frontal nonmemory functions have a predictive value for the conversion from mild cognitive impairment (MCI) to AD. Howieson et al. [47] identified TMTB as one of the most significant predictors for conversion from MCI to AD. In a recent study, executive functions such as measured by TMTB showed predictive accuracy comparable to that of MRI and cerebrospinal fluid biomarkers for the development of AD in amnesic MCI [48]. In types of dementia other than AD, impairment in TMTB, psychomotor speed and verbal fluency even dominate the clinical profile of patients, for example, in VaD and MD [49] as well as in patients with vascular white matter lesions. Impaired verbal fluency [19, 50] and reduced speed of processing on the TMT [51] may be caused by subcortical or frontal lesions, the same functions are known to be affected in patients with small vessel disease, which is related to progressive cognitive impairment and a considerable risk of developing dementia [52].

Our findings revealed that the inclusion of executive and frontal function tests as well as the survey of psychomotor speed in neuropsychological test batteries is worthwhile to improve diagnostic accuracy significantly. This finding is based on a large, reliably diagnosed sample, which is a strength of the present study; an MMSE of ≥ 24 in the patient sample ensures an early stage of the disease, and the use of Z-scores, comparable age and education within subjects provides a reliable sample that allows a high percentage of correct classification.

We also demonstrated a method for substituting missing values, which is a common issue in clinical research. Missing data are frequent in AD trials due to the age of participants and the nature of the disease. This problem can lead to bias and decreased statistical power. Analyses of the diagnostic patterns at 30 Alzheimer's Disease Centers in the United States [53] aimed to identify the most important variables for diagnosis in more than 8,000 patients. Missing values of TMTB were observed for a large number of subjects and resulted in exclusion of the test from the model in order to include as many cases as possible. Interestingly, when using the multiple imputation approach, TMTB was amongst the important predictors of diagnosed AD in that study. This finding encourages an appropriate handling of missing values. Coley et al. [54] compared current methods and attested that the Complete Case Method as well as the Last-Observation-Carried-Forward (LOCF) method were unsuitable. Multiple imputation, longitudinal mixed effects models and Z-score LOCF were declared to be more appropriate, but a careful consideration of the validity of the underlying hypotheses of methods used was advised for handling missing values. In the present study, we were confronted with a large number of missing values in the TMTB in the group of demented patients (51.8% AD and 44.6% OD). In contrast, only 1.1% of NC did not complete the test. This fact alone already confirms the relevance of TMTB, and therefore, an appropriate handling of missing values is crucial. We abstained from multiple imputations since this procedure creates multiple random results for a missing value, and we targeted instead a distinct value for every individual. Hence, we chose the procedure of deducting a penalty from the individual Z-scores of the noncompleters at the lowest possible value in order to ensure

an individual, demographically adjusted penalty for noncompliance. This method allows clinical application and shows a high prediction ability.

There are several limitations of the current study that merit further discussion. First, a lack of a gold standard in diagnosing the study sample can be mentioned, and definitive diagnoses are possible only by histopathology or autopsy. Since histopathology is rarely available, clinical diagnosis in accordance with MRI, full blood and serum analysis was used as a surrogate gold standard. Another caveat is the selection bias. A homogeneous sample like ours might be characterized by a smaller standard deviation of the predictor variable than in heterogeneous samples, higher sensitivity and specificity resulting in a facilitated differentiation of patient and control group. Additionally, the setting (Memory Clinic Basel) where the patient is examined and diagnosed is a highly specialized unit at a university hospital setting. This setting might limit the generalization of the findings to the general population in terms of background characteristics because patients at a university hospital are a priori selected by means of the referral procedure. Further, the diagnostic accuracy of both CERAD-NAB and CERAD-NAB Plus may be partially overestimated: the neuropsychological data used to diagnose patients were available to clinicians during the diagnostic process. When using the same dataset for selection and selective analysis, there may be a risk of circularity, which could confound the estimate of sensitivity and specificity and inflate the CCR. This partial circularity may have artificially heightened the diagnostic accuracy. As mentioned in the methods section, we acknowledge this potential flaw and believe to have minimized its extent through following arrangements. Assessment included not only neuropsychological examination but also a variety of specific branches (neurology, psychiatry, imaging techniques and laboratory), all of them equally engaged in establishing diagnosis. The impact of CERAD-NAB Plus variables was beyond that controlled by presentation of the neuropsychological data categorized by domains: none of the obtained cognitive domains was solely represented by a single CERAD-NAB-score.

The CERAD-NAB Plus did not significantly enhance the differential diagnosis of dementia in our study. Possible explanations for this poor discrimination might be found in the nature of information observed in this study, namely neuropsychological tests. In types of dementia other than AD, behavioral changes might be the earliest observable indicators of disease and an alternative assessment might be the method of choice in a sample like ours, consisting of patients in a very early stage of disease. For example, in FTD, inappropriate social conduct, apathy, disinhibition, perseverative behavior, loss of insight, hyperorality, and decreased speech output [55] are more easily observable than neuropsychological deficits at the onset of the disease [30]. An assessment of personality and behavioral changes might be a more promising strategy than a comparison of neuropsychological measures alone when distinguishing between patients with early-stage FTD and patients with AD. Similarly, VaD might be more reliably distinguished from AD by neuroimaging instruments rather than neuropsychological tests alone. For future research, we therefore recommend the extension of neuropsychological test batteries with measures such as questionnaires about personality and behavioral changes as well as mood changes to determine diagnostic algorithms that combine, amongst others, neuropsychological tests, imaging as well as cerebrospinal fluid analyses for better diagnosis.

In summary, our findings document both CERAD-NAB and CERAD-NAB Plus to be reliable measures when differentiating healthy individuals from patients suffering from any kind of dementia, whereas none of the two batteries seem to be the method of choice when trying to distinguish between early stages of cortical and subcortical or frontal accentuated forms of dementia. We support the notion that it is worthwhile to extend the basic CERAD-NAB by those three tests, all of them quickly to administer, in order to more reliably classify demented patients in clinical practice without overstraining patients.

Acknowledgements

We thank Michael A. Sugarman and John L. Woodard and the Research and Editing Consulting Program (RECP) for their editorial help. This work was supported by grants from the Swiss National Science Foundation (3200-049107), the Novartis Foundation, and the Velux Foundation.

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

- Schmand B, Eikelenboom P, van Gool WA: Value of neuropsychological tests, neuroimaging, and biomarkers for diagnosing Alzheimer's disease in younger and older age cohorts. *J Am Geriatr Soc* 2011;59:1705–1710.
- Salmon DP, Bondi MW: Neuropsychological assessment of dementia. *Annu Rev Psychol* 2009;60:257–282.
- Morris JC, Mohs RC, Rogers H, Fillenbaum G, Heyman A: Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacol Bull* 1988;24:641–652.
- Alzheimer's Association: World Alzheimer's Report 2010: the global economic impact of dementia. London, Alzheimer's Disease International, 2011.
- Sorbi S, Hort J, Erkinjuntti T, Fladby T, Gainotti G, Gurvit H, Nacmias B, Pasquier F, Popescu BO, Rektorova I, Religa D, Rusina R, Rossor M, Schmidt R, Stefanova E, Warren JD, Scheltens P: EFNS-ENS Guidelines on the diagnosis and management of disorders associated with dementia. *Eur J Neurol* 2012;19:1159–1179.
- Reitz C, Brayne C, Mayeux R: Epidemiology of Alzheimer disease. *Nat Rev Neurol* 2011;7:137–152.
- Ashford JW: Screening for memory disorders, dementia and Alzheimer's disease. *Aging Health* 2008;4:399–432.
- Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E: Alzheimer's disease. *Lancet* 2011;1:1019–1031.
- Twamley EW, Legendre SA, Bondi MW: Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *J Int Neuropsychol Soc* 2006;12:707–735.
- Erkinjuntti T, Gauthier S: The concept of vascular cognitive impairment. *Front Neurol Neurosci* 2009;24:79–85.
- Erkinjuntti T, Inzitari D, Pantoni L, Wallin A, Scheltens P, Rockwood K, Roman GC, Chui H, Desmond DW: Research criteria for subcortical vascular dementia in clinical trials. *J Neural Transm Suppl* 2000;59:23–30.
- Desmond DW: The neuropsychology of vascular cognitive impairment: is there a specific cognitive deficit? *J Neurol Sci* 2004;226:3–7.
- Graham NL, Emery T, Hodges JR: Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *J Neurol Neurosurg Psychiatry* 2004;75:61–71.
- Reed BR, Mungas DM, Kramer JH, Ellis W, Vinters HV, Zarow C, Jagust WC, Chui HC: Profiles of neuropsychological impairment in autopsy-defined Alzheimer's disease and cerebrovascular disease. *Brain* 2007;130:731–739.
- Graff-Radford NR, Woodruff BK: Frontotemporal dementia. *Semin Neurol* 2007;27:48–57.
- Kertesz A: Progress in clinical neurosciences: frontotemporal dementia-Pick's disease. *Can J Neurol Sci* 2006;33:141–148.
- Caso F, Corsi M, Magnani G, Fanelli G, Falautano M, Comi G, Leocani L, Minicucci F: Quantitative EEG and LORETA: valuable tools in discerning FTD from AD? *Neurobiol Aging* 2012;33:2343–2356.
- Rascovsky K, Salmon DP, Ho GJ, Galasko D, Peavy GM, Hansen LA, Thal LJ: Cognitive profiles differ in autopsy-confirmed frontotemporal dementia and AD. *Neurology* 2002;58:1801–1818.
- Diehl J, Monsch AU, Aebi C, Wagenpfeil S, Krapp S, Grimmer T, Seely W, Foerstl H, Kurz A: Frontotemporal dementia, semantic dementia, and Alzheimer's disease: the contribution of standard neuropsychological tests to differential diagnosis. *J Geriatr Psychiatry Neurol* 2005;18:39–44.
- Rascovsky K, Salmon DP, Hansen LA, Thal LJ, Galasko D: Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology* 2007;21:20–30.
- Reitan R: Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills* 1958;8:271–276.
- Benton AL, Hamsher K: Multilingual Aphasia Examination, ed 2. Iowa City, AJA Associates, 1989.
- Monsch AU, Thalmann B, Schneitter M, Bernasconi F, Aebi C, Camachova Davet Z, Stahelin HB: The Basel study on the elderly's search for preclinical cognitive markers of Alzheimer's disease. *Neurobiol Aging* 2000;21:31.
- Ehrensperger MM, Berres M, Taylor KI, Monsch AU: Early detection of Alzheimer's disease with a total score of the German CERAD. *J Int Neuropsychol Soc* 2010;16:910–920.
- American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 4. Washington DC, American Psychiatric Association, 1994.

- 26 Folstein MF, Folstein SE, McHugh PR: 'Mini Mental State' – a practical method for grading the cognitive state of patients for the clinician. *J Psych Res* 1975;12:189–198.
- 27 McKhann G, Drachman 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.
- 28 Monsch AU, Kressig RW: Specific care program for the older adults: memory clinics. *Eur Geriatr Med* 2010;1:128–131.
- 29 Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, Amaducci L, Orgogozo JM, Brun A, Hofman A, et al: Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 1993;43:250–260.
- 30 Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, Freedman M, Kertesz A, Robert PH, Albert M, Boone K, Miller BL, Cummings J, Benson DF: Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 1998;51:1546–1554.
- 31 Berres M, Monsch AU, Bernasconi F, Thalmann B, Stähelin HB: Normal ranges of neuropsychological tests for the diagnosis of Alzheimer's disease. *Stud Health Technol Inform* 2000;77:195–199.
- 32 Berres M, Zehnder A, Bläsi S, Monsch AU: Evaluation of diagnostic scores with adjustment for covariates. *Stat Med* 2008;27:1777–1790.
- 33 DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.
- 34 Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M: pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77.
- 35 Isaacs B, Kennie AT: The Set test as an aid to the detection of dementia in old people. *Br J Psychiatry* 1973;123:467–470.
- 36 Kaplan EF, Goodglass H, Weintraub S: The Boston Naming Test. Boston, Veterans Administration Medical Center, 1978.
- 37 Atkinson RC, Shiffrin RM: The control of short-term memory. *Sci Am* 1976;225:82–90.
- 38 Rosen WG, Mohs RC, Davis KL: A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984;141:1356–1364.
- 39 Masur DM, Sliwinski M, Lipton RB, Blau AD, Crystal HA: Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology* 1994;44:1427–1432.
- 40 Tierney MC, Yao C, Kiss A, McDowell I: Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology* 2005;64:1853–1859.
- 41 Amieva H, Jacqmin-Gadda H, Orgogozo J-M, Le Carret N, Helmer C, Letenneur L, Barberger-Gateau P, Fabrigoule C, Dartigues JF: The 9-year cognitive decline before dementia of the AD-type: a prospective population-based study. *Brain* 2005;128:1093–1101.
- 42 Amieva H, Le Goff M, Millet X, Orgogozo J-M, Pérès K, Barberger-Gateau P, Jacqmin-Gadda H, Dartigues JF: Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. *Ann Neurol* 2008;64:492–498.
- 43 Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C: Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *J Int Neuropsychol Soc* 2008;14:266–278.
- 44 Saxton J, Lopez OL, Ratcliff G, Dulberg C, Fried L-P, Carlson MC, Newman AB, Kuller L: Preclinical Alzheimer disease: neuropsychological test performance 1.5 to 8 years prior to onset. *Neurology* 2004;63:2341–2347.
- 45 Chen P, Ratcliff G, Belle SH, Cauley JA, DeKosky ST, Ganguli M: Patterns of cognitive decline in presymptomatic Alzheimer disease: a prospective community study. *Arch Gen Psych* 2001;58:853–858.
- 46 Bäckman L, Jones S, Berger A-K, Jonsson Laukka E, Small BJ: Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. *Neuropsychology* 2005;19:520–531.
- 47 Howieson DB, Carlsson NE, Moore MM, Waserman D, Abendroth CD, Payne-Murphy J, Kaye JA: Trajectory of mild cognitive impairment onset. *J Int Neuropsychol Soc* 2008;14:192–198.
- 48 Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack JR, Feldman HH, Bokde AL, Alexander GE, Scheltens P, Vellas B, Dubois B, Weiner M, Hampel H: Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging* 2012;33:1203–1214.
- 49 Sousa Alves G, de Oliveira Alves CE, de Oliveira Lanna ME, Ericeira-Valente L, Kenji Sudo F, Moreira D, Engelhardt E, Laks J: Clinical characteristics in subcortical ischemic white matter disease. *Arq Neuropsiquiatr* 2009;67:173–178.
- 50 Perri R, Koch G, Carlesimo GA, et al: Alzheimer's disease and frontal variant of frontotemporal dementia – a very brief battery for cognitive and behavioural distinction. *J Neurol* 2005;252:1238–1244.
- 51 Ramos-Estebanez C, Moral-Arce I, Munoz-Arrondo R, et al: Vascular cognitive impairment: prodromal stages of ischemic vascular dementia. *Dement Geriatr Cogn Disord* 2008;25:451–460.
- 52 Jokinen H, Kalska H, Ylikoski R, Madureira S, Verdelho A, van der Flier WM, Scheltens P, Barkhof F, Visser MC, Fazekas F, Schmidt R, O'Brien J, Waldemar G, Wallin A, Chabriat H, Pantoni L, Inzitari D, Erkinjuntti T: Longitudinal cognitive decline in subcortical ischemic vascular disease – the LADIS Study. *Cerebrovasc Dis* 2009;27:384–391.
- 53 Steenland K, Macneil J, Bartell S, Lah J: Analyses of diagnostic patterns at 30 Alzheimer's disease centers in the US. *Neuroepidemiology* 2010;35:19–27.
- 54 Coley N, Gardette V, Cantet C, Gillette-Guyonnet S, Nourhashemi F, Vellas B, Andrieu S: How should we deal with missing data in clinical trials involving Alzheimer's disease patients? *Curr Alz Res* 2011;8:421–433.
- 55 Neary D: Frontotemporal dementia; in Burns A, O'Brien J, Ames D (eds): *Dementia*, ed 3. London, Hodder Arnold, 2005, p 667.