Journal of Alzheimer's Disease xx (2020) x–xx
DOI 10.3233/JAD-200710
IOS Press

Value of Neuropsychological Tests to Identify Patients with Depressive Symptoms on the Alzheimer's Disease Continuum

Felix Menne[a,1,*], Carola Gertrud Schipke[b,1], Arne Klostermann[c,2], Manuel Fuentes-Casañ[a,3], Silka Dawn Freiesleben[b], Chris Bauer[d] and Oliver Peters[a,b,c]
[a]*German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany*
[b]*Charité – Universitätsmedizin Berlin, Experimental and Clinical Research Center, Berlin, Germany*
[c]*Charité – Universitätsmedizin Berlin, Department of Psychiatry and Psychotherapy, CBF, Berlin, Germany*
[d]*MicroDiscovery GmbH, Berlin, Germany*

Handling Associate Editor: Yuping Ning

Accepted 31 August 2020

Abstract.
Background: Depressive symptoms often co-occur with Alzheimer's disease (AD) and can impact neuropsychological test results. In early stages of AD, disentangling cognitive impairments due to depression from those due to neurodegeneration often poses a challenge.
Objective: We aimed to identify neuropsychological tests able to detect AD-typical pathology while taking into account varying degrees of depressive symptoms.
Methods: A battery of neuropsychological tests (CERAD-NP) and the Geriatric Depression Scale (GDS) were assessed, and cerebrospinal fluid (CSF) biomarkers were obtained. After stratifying patients into CSF positive or negative and into low, moderate, or high GDS score groups, sensitivity and specificity and area under the curve (AUC) were calculated for each subtest.
Results: 497 participants were included in the analyses. In patients with low GDS scores ($\leq$10), the highest AUC (0.72) was achieved by Mini-Mental State Examination, followed by Constructional Praxis Recall and Wordlist Total Recall (AUC = 0.714, both). In patients with moderate (11–20) and high ($\geq$21) GDS scores, Trail Making Test-B (TMT-B) revealed the highest AUCs with 0.77 and 0.82, respectively.
Conclusion: Neuropsychological tests showing AD-typical pathology in participants with low GDS scores are in-line with previous results. In patients with higher GDS scores, TMT-B showed the best discrimination. This indicates the need to focus on executive function rather than on memory task results in depressed patients to explore a risk for AD.

Keywords: Alzheimer's disease, cerebrospinal fluid, depression, executive function, memory, neuropsychology

INTRODUCTION

Cognitive impairments in old age may occur as the core symptoms of early dementia due to Alzheimer's disease (AD) [1], or they may accompany an episode of major depression (MDE) [2]. Currently, various hypotheses aiming to clarify the interrelation

[1]Current address: Predemtec AG, Berlin, Germany.
[2]Current address: Kollwitzstraße 37, Berlin, Germany.
[3]Current: Caritas-Klinik Dominikus, Berlin, Germany.
*Correspondence to: Felix Menne, Dipl.-Psych, Charité – Universitätsmedizin Berlin, Experimental and Clinical Research Center, Lindenberger Weg 80, 13125 Berlin, Germany. Tel.: +49 30 403 6688 82; E-mail: menne@predemtecdx.com.

between depression and AD exist. For example, having a history of depression has been found to increase one's risk of developing AD [3]. Depression in old age has also been suggested to represent a prodromal stage of AD rather than a risk factor for AD [1, 4, 5]. In clinical practice with geriatric patients, depressive symptoms and cognitive impairments often co-occur [6]. This makes it difficult to differentiate whether cognitive impairments are caused by depression or whether they manifest as part of a syndrome caused by AD.

There are various methods available to obtain evidence for an underlying AD pathology in cognitively impaired patients, the more biomarkers and further clinical information are available to be combined, the more accurate the diagnosis [7]. Different kinds of biomarkers help to identify the neuropathological substrates and etiology of cognitive impairments. An important source of information in the diagnosis of AD are brain magnetic resonance imaging (MRI) scans [8]. However, MRI scans can be contraindicated due to pacemakers or other electrical implants, anxiety, or economic reasons.

The quantification of total-Tau (*t*-Tau) and amyloid-β 1-42 ($A\beta_{42}$) proteins in the cerebrospinal fluid (CSF) have been established to detect early AD-typical pathology with high sensitivity and specificity [9]. Specialized memory clinics may recommend the quantification of CSF biomarkers in late-life depression to help determine whether an underlying AD pathology exists [10]. However, lumbar punctures for obtaining CSF might be perceived as highly invasive by some patients. Furthermore, processing and analyzing CSF is highly demanding and alterations in sample processing can lead to varying results [11]. Lastly, lumbar punctures might be contraindicated in some patients taking anticoagulants or suffering from conditions like scoliosis.

Neuropsychological assessments together with clinical information is the basis to determine different stages of cognitive decline [12, 13]. Ideally, as much additional diagnostic evidence as possible should be used to accurately diagnose AD [7]. However, for different reasons mentioned earlier, some methods might not be available. Thus, identifying easy to conduct, sensitive, and valid neuropsychological tests can add to a more accurate diagnosis of underlying AD pathology.

Previous studies have aimed to identify neuropsychological tests able to differentiate between early AD and late-onset depression. There is evidence that the meaningfulness of psychological test results can differ depending on the affective state of patients [14, 15], which should be considered when interpreting test results. A frequently cited test helpful in distinguishing cognitive impairments due to depression from those due to AD is the Clock Drawing Test (CDT [16]), although contradictory findings exist regarding the extent to which the CDT is able to fulfill this task [17, 18]. When examining episodic memory function, both patients with early AD and MDE show a below-average performance on immediate and delayed recall tests. However, depressed patients retain the learned information better than early AD patients, as measured by recognition tasks [19]. Moreover, there is evidence that depression in AD patients additionally impairs performance in executive function tasks as measured by the Trail Making Test Part B (TMT-B) compared to AD patients without depression [20].

Many publications address the differences in cognitive domains between depression and AD. However, depression is not black or white, but rather there are varying stages in affective mood between clinically depressed and non-depressed. Taking these considerations into account, we aimed to examine the effect of varying numbers of depressive symptoms when interpreting neuropsychological results. In our approach, we wanted to examine the value of different neuropsychological tests to detect AD-typical pathology in old age CSF classified patients. We hypothesized that depending on the number of depressive symptoms patients present, neuropsychological tests would vary in their ability to differentiate between patients with and without AD-typical changes in CSF.

METHODS

Participants

The sample consisted of memory clinic patients presenting with subjective cognitive impairment between 2007 and 2018. Routine clinical practice comprised of a medical case history assessment, psychopathological examination, comprehensive neuropsychological testing, cranial imaging, and a lumbar puncture to assess CSF biomarkers ($A\beta_{40}$, $A\beta_{42}$, and *t*-Tau). DSM-IV/-V diagnosis for each patient was reached by a consensus panel. The study was conducted in accordance with the Declaration of Helsinki. Ethical vote was obtained from the Ethics Committee of the Charité Universitätsmedizin Berlin, vote number EA4/057/20.

Neuropsychological tests and clinical scales

We assessed cognitive performance with the Consortium to establish a registry for Alzheimer's disease neuropsychological test battery (CERAD-NP). CERAD-NP is a standardized instrument used in routine clinical practice to assess and stage AD-typical cognitive impairments [13].

Specifically, the CERAD-NP consists of the Mini-Mental State Examination (MMSE) [21], phonemic fluency, and visual naming (Boston Naming Test [22]), tests for constructional praxis and constructional praxis delayed recall and verbal memory tasks. Furthermore, tests to measure processing speed and executive function, namely the Trail Making Test A (TMT-A) and the TMT-B [23], as well as the CDT [16] were performed. Results on each CERAD-NP-subscale are z-standardized, taking gender, age, and years of education into account.

Depressive symptoms were assessed with the original 30-item version of the Geriatric Depression Scale (GDS; yes/no dichotomous scale, range 0–30, scores proportional to depressive symptoms) [24]. The GDS is a self-administered questionnaire shown to be a valid instrument to help identify late-life depression [25]. A cut-off score of $\geq$11 can be seen as a possible indicator of depression, as it has been shown to have a sensitivity of 84% and a specificity of 95% for accurately detecting late-life depression [26]. According to our clinical experience, GDS scores of 21 or higher are highly indicative of clinical depression. For these reasons, we decided to divide patients into one of three GDS subgroups, namely patients with a GDS-score $\leq$10 (low GDS), 11–20 (moderate GDS), and $\geq$21 (high GDS).

Cerebrospinal fluid

CSF was collected and analyzed according to a standardized protocol described in detail elsewhere [27]. As it is known that differing and analytical procedures and lot-to-lot variation of analytical kits can strongly influence CSF biomarkers [28], we established the following CSF biomarker cut-offs as indicative of AD-typical pathology in our memory clinic: $A\beta_{42} \leq 600$ pg/ml (sensitivity 0.82, specificity 0.80) or ratio $A\beta_{42}/A\beta_{40} \leq 0.065$ (sensitivity 0.80, specificity 0.75), added by *t*-Tau $\geq$ 350 pg/ml (sensitivity 0.74, specificity 0.78).

Following the NIA-AA research framework [29], we defined CSF positive patients showing both amyloid-pathology (*A*+) and neurodegeneration (*N*+). For the analyses presented here, we defined patients as having AD-typical pathology (i.e., CSF-positive) when *t*-Tau $\geq$350 pg/ml and $A\beta_{42}$ $\leq$ 600 pg/ml. In CSF negative patients, the cut-offs were *t*-Tau < 350 pg/ml and $A\beta_{42} > 600$ pg/ml, corresponding to A- and N-.

Inclusion and exclusion criteria

We included patients that underwent a complete diagnostic assessment in our memory clinic as described above and who had an MMSE score of 24 or higher with the aim to identify patients with mild or no objective cognitive deficits.

We excluded patients that did not fulfill our established CSF positive or CSF negative criteria. No further exclusion criteria (e.g., diagnosis or medication) were defined in order to better reflect a cohort of patients clinicians face in their everyday work.

Statistical analyses

Data were analyzed using the statistical software "R", version 3.2.4. The authors were blind to patients' diagnosis.

After dividing patients into CSF positive and CSF negative, a single value classification was performed. We calculated receiver operating characteristic (ROC) curves for all neuropsychological tests by calculating the sensitivity and specificity for each value of the neuropsychological test results. The performance of the classification was assessed using the area under the curve (AUC). The AUC typically ranges between 0.5 and 1, with an AUC of 1 indicating perfect discrimination and an AUC of 0.5 reflecting a random classification. Confidence intervals and *p*-values to compare ROC curves were calculated according to the Delong algorithm.

For further analyses, we formed the cognitive domains Recall (Wordlist Recall, Constructional Praxis Savings, Discriminability) and Executive Function (Semantic Fluency, Trail Making Test A and B) and calculated AUC values as described above.

To investigate the relation between classification performance and depressive symptoms, we performed a series of single value classifications for patients with increasing GDS scores. For a given GDS score, we selected all patients with a score of $\pm$10 and performed the single value classification as described above.

For the descriptive statistics, Student's *t*-tests or when appropriate non-parametric Wilcoxon

Table 1
Demographics, clinical scale scores and CSF data of CSF-negatives and -positives

	CSF-positive	CSF-negative	*p*
	190	307	
Female sex (%)	53	44	0.15
Years of education	13.4 ± 3.0	13.6 ± 2.9	0.61
Age	68.0 ± 9.0	69.8 ± 9.9	0.61
t-Tau (pg/ml)	549 ± 284	228 ± 65	**<0.001**
$A\beta_{42}$ (pg/ml)	391 ± 115	1054 ± 339	**<0.001**
Ratio *t*-Tau/$A\beta_{42}$	1.52 ± 0.96	0.24 ± 0.10	**<0.001**
MMSE	26.4 ± 1.7	27.7 ± 1.6	**<0.001**
Mean GDS subgroup 0–10 ($n = 214$)	5.8 ± 2.7	6.0 ± 2.9	0.8
Mean GDS subgroup 11–20 ($n = 197$)	14.7 ± 2.8	14.9 ± 2.9	1.0
Mean GDS subgroup 21–30 ($n = 86$)	23.7 ± 2.6	24.1 ± 2.6	1.0

two-sample tests were used to investigate differences between group means on continuous variables.

RESULTS

Patient selection

A total of 2,101 patients underwent a complete diagnostic assessment at our memory clinic between 2007 and 2018. A total of 1,414 patients with an MMSE score of < 24 were excluded from further analyses.

Patient characteristics

Of the remaining 687 patients, 190 had CSF biomarker results that did not fulfill criteria for either being CSF positive (*A + and N+*) or CSF negative (A- *and* N-) and were excluded from further analyses. Of the remaining 497 patients, 307 were defined as being CSF negative and 190 as CSF positive. Table 1 provides information on patients' demographics, MMSE, GDS, and CSF data.

The 190 CSF positive patients performed significantly worse in all CERAD-NP subtests than CSF negative patients. CSF positive patients scored lower than –1.5 SD below the mean in Constructional Praxis Recall (–1.8 ± 1.3), World List Trial 3 (–1.7 ± 1.5), Wordlist Recall (–1.7 ± 1.4), and Wordlist Total (–1.8 ± 1.5) tests. CSF negative patients yielded *z*-scores ≥–1.5 SD in all CERAD-NP subtests, indicating normative cognitive performance. Table 2 shows the complete list of neuropsychological test performance by CSF group.

Table 2
Neuropsychological test performance

	CSF-positive	CSF-negative	*p*
BNT	–0.1 ± 1.3	0.2 ± 1.2	**<0.001**
CDT	2.2 ± 1.0	1.6 ± 0.9	**<0.001**
CP	–0.2 ± 1.3	0.1 ± 1.2	**<0.05**
CPR	–1.8 ± 1.3	–0.4 ± 1.6	**<0.001**
CPS	–1.5 ± 1.3	–0.4 ± 1.2	**<0.001**
MMSE	26.4 ± 1.7	27.7 ± 1.6	**<0.001**
SF	–0.9 ± 1.1	–0.4 ± 1.3	**<0.001**
TMT-A	–0.8 ± 1.3	0 ± 1.4	**<0.001**
TMT-B	–1.2 ± 1.3	0 ± 1.7	**<0.001**
TMT-B/A	–0.6 ± 1.1	–0.1 ± 1.3	**<0.001**
WL_discr	–1.3 ± 1.4	–0.5 ± 1.4	**<0.001**
WL_I	–0.8 ± 1.3	–0.2 ± 1.1	**<0.001**
WL_R	–1.7 ± 1.4	–0.6 ± 1.2	**<0.001**
WL_S	–1.4 ± 2.2	–0.4 ± 1.8	**<0.001**
WL_total	–1.8 ± 1.5	–0.7 ± 1.4	**<0.001**
WL1	–1.2 ± 1.2	–0.5 ± 1.2	**<0.001**
WL2	–1.4 ± 1.3	–0.6 ± 1.3	**<0.001**
WL3	–1.7 ± 1.5	–0.6 ± 1.4	**<0.001**

MMSE and CDT mean raw scores as well as CERAD-NP mean z-standardized scores in the groups of CSF-positives and -negatives (sorted alphabetically). BNT, Boston Naming Test; CDT, Clock Drawing Test. CP: Constructional Praxis; CPR, Constructional Praxis Recall; CPS, Constructional Praxis Savings; MMSE, Mini-Mental Status Examination; SF, Semantic Fluency; TMT-A, Trail-Making Test A; TMT-B, Trail-Making Test B; TMT-B/A, Ratio of TMT B/A; WL_discr, Wordlist Discrimination; WL_I, Wordlist Intrusions; WL_R, Wordlist Delayed Recall; WL_S, Wordlist Savings; WL_total, Wordlist Total of immediately recalled words; WL_1, Wordlist 1st trial; WL_2, Wordlist 2nd trial; WL_3, Wordlist 3rd trial. All group differences showed significance.

Patient characteristics by GDS subgroup

In patients with low GDS scores (≤10, $n = 214$), 102 were CSF positive (47%). In those with moderate GDS scores (11–20, $n = 197$), 73 were CSF positive (37%), and in those with high GDS scores (≥21, $n = 86$), 15 were CSF positive (17%).

Discrimination accuracy of neuropsychological tests between CSF groups and GDS subgroups

In patients with GDS scores ≤10, the neuropsychological tests with the highest specificity and sensitivity in differentiating between CSF positive and CSF negative were the MMSE (AUC = 0.72), Constructional Praxis Recall (0.71), and Wordlist Total (0.71). In patients with GDS scores between 11–20, the Trail Making Test-B (0.77), Wordlist Discriminability (0.75), and Wordlist Recall (0.75) showed the highest specificity and sensitivity. The neuropsychological tests with the highest specificity and sensitivity to differentiate between CSF groups

Table 3
AUC of neuropsychological tests

	AUC all	AUC GDS ≤10	AUC GDS 11–20	AUC GDS 21–30
WL_R	0.737 (305/182) [0.69,0.784]	0.706 (112/99) [0.634,0.777]	0.752 (122/69) [0.678,0.826]	0.783 (71/14) [0.659,0.906]
CPR	0.732 (307/190) [0.687,0.778]	0.714 (112/102) [0.644,0.783]	0.733 (124/73) [0.662,0.804]	0.776 (71/15) [0.639,0.912]
CPS	0.728 (306/189) [0.681,0.775]	0.706 (111/102) [0.635,0.777]	0.738 (124/72) [0.666,0.81]	0.733 (71/15) [0.573,0.893]
WL_total	0.719 (307/190) [0.673,0.765]	0.714 (112/102) [0.645,0.783]	0.723 (124/73) [0.648,0.798]	0.735 (71/15) [0.587,0.884]
WL_3	0.719 (307/190) [0.672,0.766]	0.705 (112/102) [0.635,0.775]	0.737 (124/73) [0.662,0.811]	0.721 (71/15) [0.565,0.876]
MMSE	0.713 (307/190) [0.667,0.758]	0.72 (112/102) [0.653,0.788]	0.68 (124/73) [0.603,0.756]	0.761 (71/15) [0.634,0.887]
TMT-B	0.708 (307/190) [0.663,0.754]	0.643 (112/102) [0.569,0.717]	0.766 (124/73) [0.7,0.831]	0.816 (71/15) [0.704,0.928]
WL_2	0.68 (307/190) [0.632,0.728]	0.667 (112/102) [0.596,0.739]	0.68 (124/73) [0.601,0.759]	0.751 (71/15) [0.603,0.898]
WL_discr	0.677 (307/190) [0.629,0.726]	0.645 (112/102) [0.571,0.72]	0.753 (124/73) [0.684,0.823]	0.592 (71/15) [0.447,0.738]
WL_1	0.673 (307/190) [0.625,0.721]	0.678 (112/102) [0.607,0.749]	0.668 (124/73) [0.589,0.747]	0.671 (71/15) [0.508,0.834]
WL_sav	0.669 (304/182) [0.616,0.721]	0.638 (111/99) [0.561,0.715]	0.683 (122/69) [0.597,0.768]	0.752 (71/14) [0.62,0.883]
CDT	0.664 (307/190) [0.618,0.71]	0.633 (112/102) [0.563,0.703]	0.665 (124/73) [0.59,0.739]	0.786 (71/15) [0.67,0.902]
TMT-A	0.639 (307/190) [0.589,0.688]	0.634 (112/102) [0.56,0.708]	0.702 (124/73) [0.628,0.776]	0.619 (71/15) [0.434,0.804]
TMT-B/A	0.63 (307/190) [0.581,0.68]	0.552 (112/102) [0.475,0.63]	0.657 (124/73) [0.581,0.734]	0.762 (71/15) [0.626,0.897]
WL_I	0.626 (307/190) [0.574,0.678]	0.627 (112/102) [0.552,0.703]	0.635 (124/73) [0.55,0.719]	0.521 (71/15) [0.352,0.69]
SF	0.622 (302/189) [0.572,0.672]	0.618 (112/101) [0.542,0.694]	0.623 (120/73) [0.544,0.701]	0.641 (70/15) [0.463,0.82]
BNT	0.592 (307/190) [0.54,0.644]	0.553 (112/102) [0.475,0.632]	0.6 (124/73) [0.519,0.682]	0.752 (71/15) [0.608,0.895]
CP	0.556 (307/189) [0.502,0.61]	0.521 (112/102) [0.442,0.6]	0.581 (124/72) [0.493,0.668]	0.581 (71/15) [0.371,0.79]

Area under the curve (AUC) as well as number of subjects (CSF negative/positive) and [confidence interval] of each neuropsychological test in respective GDS score groups and irrespective of GDS score. WL_1, Wordlist 1st trial; CPR, Constructional Praxis Recall; CPS, Constructional Praxis Savings; WL_3, Wordlist 3rd trial; WL_R, Wordlist Delayed Recall; MMSE, Mini-Mental Status Examination; TMT-A, Trail-Making Test A; WL_I, Wordlist Intrusions; WL_S, Wordlist Savings; WL_discr, Wordlist Discrimination; WL_2, Wordlist 2nd trial; CDT, Clock Drawing Test; SF, Semantic Fluency; TMT-B, Trail-Making Test B; WL_total, Wordlist Total of immediately recalled words; TMT-B/A, Ratio of TMT B/A; BNT, Boston Naming Test; CP, Constructional Praxis.

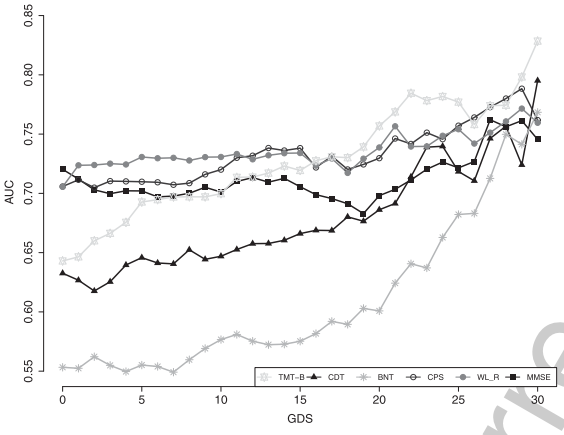

Fig. 1. AUC of selected neuropsychological tests with increasing GDS. TMT-B, Trail-Making Test B; CDT, Clock Drawing Test; BNT, Boston Naming Test; CPS, Constructional Praxis Savings; WL_R, Wordlist Delayed Recall; MMSE, Mini-Mental Status Examination.

with GDS scores between 21–30 were the Trail Making Test-B (0.82), CDT (0.79), and Wordlist Recall (0.78) tests. An overview of AUC values for all neuropsychological tests by GDS subgroup is presented in Table 3.

When analyzing the discriminatory power throughout GDS scores (0–30), we find a rise in the AUC values of several neuropsychological tests with increasing GDS scores. In Fig. 1, we present six CERAD-NP subtests we selected because of their significant rise in AUC with increasing GDS values. In particular, the TMT-B and especially the Boston Naming Test (BNT) test exhibit a marked rise in AUC values with higher GDS scores. There are significant differences when comparing AUCs of the TMT-B (AUC: 0.64 versus 0.82, $p < 0.02$) and the BNT (AUC: 0.55 versus 0.75, $p < 0.02$) between the two groups with low (≤10) and high (≥21) GDS scores.

No significant differences can be found when comparing the cognitive domains Recall (AUC: 0.71 versus 0.76, $p = 0.52$) and Executive Function (AUC: 0.65 versus 0.75, $p = 0.23$) between high and low GDS score groups (Supplementary Table 1).

DISCUSSION

Cognitive impairments in old age have numerous causes. Better understanding the etiology of cognitive decline is a prerequisite for appropriate treatment. Here, we explored the sensitivity and specificity of different neuropsychological tests to identify cognitive impairments typical of AD pathology in the presence of varying degrees of depressive symptoms in patients verified for AD-typical CSF biomarkers. Our

findings support our hypothesis that depending on the number of depressive symptoms, neuropsychological tests will vary in their ability to differentiate between subjects with and without AD-typical changes in CSF. We found that in subjects with a moderate to high number of depressive symptoms, assessing executive function with the TMT-B has the highest power to discriminate between CSF-positive and CSF-negative patients. Furthermore, we observed an increasing discriminatory power of several CERAD-NP subtests over the course of rising GDS scores.

Upon closer examination of different neuropsychological subtests and their ability to discriminate between CSF-positive and negative subjects, the CDT showed to be a valuable instrument in patients with high GDS scores between 21–30. However, differentiation accuracy was lower in patients with lower GDS scores. Although the CDT is largely used to assess AD-typical cognitive impairments and has shown acceptable sensitivity and specificity in patients with depression [17], its clinical value remains controversial. It has also been shown that the CDT lacks sensitivity in mildly impaired patients [18] and is not well suited to differentiate between AD patients and patients suffering from other types of dementia [30].

The MMSE is known to have limitations in detecting cognitive impairments in early AD [31], which appears to be mainly due to its ceiling and floor effects and due to the marked impact of age and education on test results [32]. Interestingly, our results showed that the MMSE had the highest power (AUC = 0.72) in distinguishing between CSF positive and CSF negative patients in the low GDS subgroup. This is most likely due to the broad range of cognitive domains that are covered by the MMSE. However, it seems that in patients with higher GDS scores, other tests outperform the MMSE.

The TMT-B test was best at differentiating between CSF positive and negative patients with moderate to high GDS scores (11–30). The TMT-B assesses, among others, executive function, which has been shown to be impaired not only in mild AD [33] but also in earlier stages of AD (i.e., MCI due to AD) [34] and there is evidence the TMT-B may help distinguish between cognitively healthy controls, AD, and depressed patients [35, 36]. Our results are inline with these previous findings. Hence, we can confirm the value of testing patients' executive function to establish a differential neuropsychological diagnosis.

Interestingly, with increasing GDS scores, we observed a broad rise in the AUC values of a few CERAD-NP subtests. It has been shown before that comorbid depression influences AD patients' test performance in the TMT-B [20]. In our data, higher depressive symptoms in CSF positive patients seem to more strongly influence test performance than in CSF negative patients. Since being at risk for AD as defined by CSF-typical biomarker changes typically leads to a significant difference in test performance compared to CSF negative patients [37], a concurrent high number of depressive symptoms might lead to an even more pronounced difference in test performance. This can be seen as a "double hit", resulting in the higher power of a few neuropsychological tests to differentiate between the two CSF groups. This might also explain the difference between the AUCs of the TMT-A and TMT-B tests. The higher cognitive demand of the TMT-B compared to the TMT-A test might lead to worse performance in CSF patients with higher GDS scores compared to CSF negative patients.

Our findings stress the differential diagnostic value of specific neuropsychological test results of old age patients presenting with depressive symptoms. Indeed, depending on the level of depressive symptoms, traditionally used tests like the MMSE and the CDT showed less power than other tests such as the TMT-B test in discriminating between patients with and without AD-typical CSF pathology. Therefore, for patients presenting to a memory clinic with suspected or clinically manifest depression, we recommend focusing on tests that assess executive function rather than the MMSE, CDT, or verbal memory tests for higher diagnostic differentiation. Doing so can help guide clinicians in their decision of whether further diagnostic measures are warranted.

We consider the high number of patients with available CSF data a strength of this analysis. To the best of our knowledge, we are not aware of any published data of monocentric databases with a similar amount of CSF data available. Furthermore, using patients' CSF data and neuropsychological test results rather than their diagnosis reduces the risk of being biased by their clinical diagnosis when interpreting our findings. Moreover, few publications regarding neuropsychological test performance in early AD patients with moderate or high depressive symptoms are available, as mood disorders are often exclusion criteria in studies on neurodegenerative disorders.

The cross-sectional nature of the study may be seen as a limitation. Since no follow-up examinations were conducted, it remains unclear whether the observed CSF abnormalities resulted in neuropsychological and GDS score changes or whether these changes were present before CSF abnormalities. Furthermore, no phosphorylated tau (*p*-Tau) data was available, which together with Aβ defines AD according to the NIA-AA research framework [29]. Moreover, since the GDS is a self-reported measure, scores might not accurately reflect the severity of depressive symptoms as would be obtained by a trained clinician. This may have under- or overestimated the actual degree of depressive symptoms in some patients, which might additionally be influenced by antidepressant or anxiolytic medication. Lastly, the unequal GDS subgroups and CSF group sizes limit statistical power, thus results presented here must be interpreted with caution. These differences are noticeable especially in the ratio of CSF positive versus negative subjects in the group of GDS scores $\geq$21. This is likely due to the fact that the majority of our patients presenting with memory concerns who have high GDS scores suffer only from depression and less likely from an additional underlying neurodegenerative process. Furthermore, we suspect that patients with high GDS scores who at the same time suffer from a neurodegenerative disease would be more severely impaired and thus have an MMSE score below 24, which we excluded in this study.

Our results support previous studies identifying neuropsychological tests that more accurately differentiate between patients with MCI, mild AD, or MDE. However, especially in mildly cognitively impaired individuals, differentiation based on neuropsychological tests alone is difficult [38, 39]. Our findings strengthen existing results regarding which neuropsychological tests used in clinical routine practice are best at differentiating between CSF positive and CSF negative patients while considering varying degrees of depressive symptoms.

ACKNOWLEDGMENTS

We want to thank all participants and their caregivers for contributing to this paper.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

FM and CGS are employees of Predemtec AG.

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/20-0710r2).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-200710.

REFERENCES

[1] Singh-Manoux A, Dugravot A, Fournier A, Abell J, Ebmeier K, Kivimäki M, Sabia S (2017) Trajectories of depressive symptoms before diagnosis of dementia: A 28-year follow-up study. *JAMA Psychiatry* **74**, 712-718.

[2] Rock PL, Roiser JP, Riedel WJ, Blackwell AD (2014) Cognitive impairment in depression: A systematic review and meta-analysis. *Psychol Med* **44**, 2029-2040.

[3] Ownby RL, Crocco E, Acevedo A, John V, Loewenstein D (2006) Depression and risk for Alzheimer disease. *Arch Gen Psychiatry* **63**, 530-538.

[4] Mirza SS, Wolters FJ, Swanson SA, Koudstaal PJ, Hofman A, Tiemeier H, Ikram MA (2016) 10-year trajectories of depressive symptoms and risk of dementia: A population-based study. *Lancet Psychiatry* **3**, 628-635.

[5] Javaherian K, Newman BM, Weng H, Hassenstab J, Xiong C, Coble D, Fagan AM, Benzinger T, Morris JC (2019) Examining the complicated relationship between depressive symptoms and cognitive impairment in preclinical Alzheimer disease. *Alzheimer Dis Assoc Disord* **33**, 15-20.

[6] Chi S, Wang C, Jiang T, Zhu X-C, Yu J-T, Tan L (2015) The prevalence of depression in Alzheimer's disease: A systematic review and meta-analysis. *Curr Alzheimer Res* **12**, 189-198.

[7] Frölich L, Peters O, Lewczuk P, Gruber O, Teipel SJ, Gertz HJ, Jahn H, Jessen F, Kurz A, Luckhaus C, Hüll M, Pantel J, Reischies FM, Schröder J, Wagner M, Rienhoff O, Wolf S, Bauer C, Schuchhardt J, Heuser I, Rüther E, Henn F, Maier W, Wiltfang J, Kornhuber J (2017) Incremental value of biomarker combinations to predict progression of mild cognitive impairment to Alzheimer's dementia. *Alzheimers Res Ther* **9**, 84.

[8] Johnson KA, Fox NC, Sperling RA, Klunk WE (2012) Brain imaging in Alzheimer disease. *Cold Spring Harb Perspect Med* **2**, a006213.

[9] Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L (2006) Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: A follow-up study. *Lancet Neurol* **5**, 228-234.

[10] Liguori C, Pierantozzi M, Chiaravalloti A, Sancesario GM, Mercuri NB, Franchini F, Schillaci O, Sancesario G (2018) When cognitive decline and depression coexist in the elderly: CSF biomarkers analysis can differentiate Alzheimer's disease from late-life depression. *Front Aging Neurosci* **10**, 38.

[11] Mattsson N, Blennow K, Zetterberg H (2010) Inter-laboratory variation in cerebrospinal fluid biomarkers for Alzheimer's disease: United we stand, divided we fall. *Clin Chem Lab Med* **48**, 603-607.

[12] Morris JC (1997) Clinical Dementia Rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatrics* **9**(Suppl 1), 173-176.

[13] Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C (1989) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* **39**, 1159-1165.

[14] Kunig G, Jager M, Stief V, Kaldune A, Urbaniok F, Endrass J (2006) The impact of the CERAD-NP on diagnosis of cognitive deficiencies in late onset depression and Alzheimer's disease. *Int J Geriatr Psychiatry* **21**, 911-916.

[15] Hofrichter NA, Dick S, Riemer TG, Schleussner C, Goerke M, Mell T, Heinz A, Rapp MA (2014) Impact of comorbid depression on serial position effects in alzheimer's disease. *GeroPsych (Bern)* **27**, 161-169.

[16] Shulman KI, Pushkar Gold D, Cohen CA, Zucchero CA (1993) Clock-drawing and dementia in the community: A longitudinal study. *Int J Geriatr Psychiatry* **8**, 487-496.

[17] Herrmann N, Kidron D, Shulman KI, Kaplan E, Binns M, Leach L, Freedman M (1998) Clock tests in depression, Alzheimer's disease, and elderly controls. *Int J Psychiatry Med* **28**, 437-447.

[18] Lee H, Swanwick GR, Coen RF, Lawlor BA (1996) Use of the clock drawing task in the diagnosis of mild and very mild Alzheimer's disease. *Int Psychogeriatr* **8**, 469-476.

[19] Wright SL, Persad C (2007) Distinguishing between depression and dementia in older persons: Neuropsychological and neuropathological correlates. *J Geriatr Psychiatry Neurol* **20**, 189-198.

[20] Nakaaki S, Murata Y, Sato J, Shinagawa Y, Tatsumi H, Hirono N, Furukawa TA (2007) Greater impairment of ability in the divided attention task is seen in Alzheimer's disease patients with depression than in those without depression. *Dement Geriatr Cogn Disord* **23**, 231-240.

[21] Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.

[22] Kaplan E, Goodglass H, Weintraub S (1983) *The Boston Naming Test*, Lea and Febiger, Philadelphia.

[23] Reitan RM (1955) The relation of the Trail Making Test to organic brain damage. *J Consult Psychol* **19**, 393-394.

[24] Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO (1982) Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* **17**, 37-49.

[25] Yesavage JA, Sheikh JI (1986) Geriatric Depression Scale (GDS). *Clin Gerontol* **5**, 165-173.

[26] Brink TL, Yesavage JA, Owen L, Heersema PH, Adey M, Rose TL (1982) Screening tests for geriatric depression. *Clin Geron* **1**, 37-43.

[27] Schipke CG, Prokop S, Heppner FL, Heuser I, Peters O (2011) Comparison of immunosorbent assays for the quantification of biomarkers for Alzheimer's disease in human cerebrospinal fluid. *Dement Geriatr Cogn Disord* **31**, 139-145.

[28] Vos SJB, Visser PJ, Verhey F, Aalten P, Knol D, Ramakers I, Scheltens P, Olde Rikkert MGM, Verbeek MM, Teunissen CE (2014) Variability of CSF Alzheimer's disease biomarkers: Implications for clinical practice. *PLoS One* **9**, e100784.

[29] Jack Jr. CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, Contributors (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535-562.

[30] Tan LPL, Herrmann N, Mainland BJ, Shulman K (2015) Can clock drawing differentiate Alzheimer's disease from other dementias? *Int Psychogeriatrics* **27**, 1649-1660.

[31] Mitchell AJ (2009) A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatr Res* **43**, 411-431.

[32] Tombaugh TN, McIntyre NJ (1992) The mini-mental state examination: A comprehensive review. *J Am Geriatr Soc* **40**, 922-935.

[33] Lafleche G, Albert MS (1995) Executive function deficits in mild Alzheimer's disease. *Neuropsychology* **9**, 313-320.

[34] Wylie SA, Ridderinkhof KR, Eckerle MK, Manning CA (2007) Inefficient response inhibition in individuals with mild cognitive impairment. *Neuropsychologia* **45**, 1408-1419.

[35] Nathan J, Wilkinson D, Stammers S, Low JL (2001) The role of tests of frontal executive function in the detection of mild dementia. *Int J Geriatr Psychiatry* **16**, 18-26.

[36] Chen P, Ratcliff G, Phil D, Belle SH, Cauley JA, Dekosky ST, Ganguli M (2000) Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. *Neurology* **55**, 1847-1853.

[37] Wagner M, Wolf S, Reischies FM, Daerr M, Wolfsgruber S, Jessen F, Popp J, Maier W, Hüll M, Frölich L, Hampel H, Perneczky R, Peters O, Jahn H, Luckhaus C, Gertz HJ, Schröder J, Pantel J, Lewczuk P, Kornhuber J, Wiltfang J (2012) Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology* **78**, 379-386.

[38] O'Carroll RE, Conway S, Ryman A, Prentice N (1997) Performance on the delayed word recall test (DWR) fails to differentiate clearly between depression and Alzheimer's disease in the elderly. *Psychol Med* **27**, 967-971.

[39] Coen RF, Kirby M, Swanwick GRJ, Maguire CP, Walsh JB, Coakley D, O'neill D, Lawlor BA (1997) Distinguishing between patients with depression or very mild Alzheimer's disease using the delayed-word-recall test. *Dement Geriatr Cogn Disord* **8**, 244-247.