

Received: 2015.07.19  
Accepted: 2015.08.14  
Published: 2015.08.31

# Selective Ability of Some CANTAB Battery Test Measures to Detect Cognitive Response to a Single Dose of Donepezil in Alzheimer Disease

Authors' Contribution:  
Study Design A  
Data Collection B  
Statistical Analysis C  
Data Interpretation D  
Manuscript Preparation E  
Literature Search F  
Funds Collection G

ABCDEF 1,2 **Jurgita Kuzmickienė**  
ABCDEF 1,2 **Gintaras Kaubrys**

1 Clinic of Neurology and Neurosurgery, Faculty of Medicine, Vilnius University, Vilnius, Lithuania  
2 Center of Neurology, Vilnius University Hospital Santariskiu Klinikos, Vilnius, Lithuania

**Corresponding Author:** Gintaras Kaubrys; e-mail: [gintaras.kaubrys@santa.lt](mailto:gintaras.kaubrys@santa.lt)  
**Source of support:** Departmental sources

**Background:** The Cambridge Neuropsychological Test Automated Battery (CANTAB) was used to explore which tests and their measures are able to detect cognitive change after a single dose of donepezil in Alzheimer disease (AD) patients. The aim of this study was to establish the ability of CANTAB tests and their measures to detect cognitive change after a single 5-mg dose of donepezil in treatment-naïve AD patients.





**Material/Methods:** We enrolled 62 treatment-naïve AD patients and 30 healthy controls in this prospective, randomized, rater-blinded study. AD patients were randomized to 2 groups: the AD+ group received donepezil after the first CANTAB testing and the AD- group remained treatment-naïve at second testing. The time period between repeated testing was 4 hours. Parallel versions of CRT, SOC, PAL, SWM, and PRM tests were used.

**Results:** All groups did not differ according to age, education, gender, or depression ( $p > 0.05$ ). AD+ and AD- groups did not differ according to MMSE. SOC, PAL, PRM, and SWM tests distinguished AD from controls. Eight measures of PAL and PRM had a strong correlation with MMSE ( $r > 0.7$ ). Repeated-measures ANOVA with Bonferroni post-hoc test showed the difference of change in AD+ and AD- groups between first and second CANTAB testing in 7 PAL measures. AD+ and AD- groups differed in the second testing by 7 PAL measures. Four PAL measures differed in first and second testing within the AD+ group.

**Conclusions:** The CANTAB PAL test measures, able to detect cognitive change after a single dose of donepezil in AD patients, are: PAL mean trials to success, total errors (adjusted), total errors (6 shapes, adjusted), and total trials (adjusted).

**MeSH Keywords:** **Alzheimer Disease • Cholinesterase Inhibitors • Dementia • Neurobehavioral Manifestations • Paired-Associate Learning • Treatment Outcome**

**Full-text PDF:** <http://www.medscimonit.com/abstract/index/idArt/895381>

 5137  4  1  33



## Background

Alzheimer disease (AD) is a neurodegenerative disease accompanied by progressive cognitive decline leading to dementia. AD accounts for up to two-thirds of cases of dementia in the elderly population [1,2]. Although the pathogenesis of AD is not fully understood, it is established that the deficit of acetylcholine in the brain has a major role in occurrence of cognitive symptoms in AD [3,4]. The cholinergic system is involved in the storage and retrieval of new information and other aspects of memory [5]. Stimulation of the cholinergic system with cholinesterase (ChE) inhibitors diminishes the degradation of acetylcholine and increases its accumulation in the synapse, thereby improving cognitive functions in AD [6,7]. Donepezil is a centrally acting reversible and selective acetylcholinesterase inhibitor. It has a long duration of inhibitory action – biological half-life of about 70 hours, 100% bioavailability, and a greater specificity for brain tissue [8]. Many studies demonstrated the efficacy of donepezil in patients with AD [6,7]. There are many cognitive tests that are used in AD to determine general cognitive function and specific cognitive deficits in specific cognitive domains. Tests and batteries most frequently used in clinical practice and clinical trials are the so-called “paper-pencil” tests. Most of clinical trials with donepezil have shown that it is possible to detect significant improvement of cognition after 4 weeks of treatment [6,7,9]. The primary cognitive assessment tool used in most trials for mild and moderate AD is the Alzheimer Disease Assessment Scale – cognitive subscale (ADAS-cog), which is a typical paper-pencil test. It might be useful to be able to identify psychometric instruments able to reliably detect cognitive change (improvement) much earlier, ideally, after a single dose of symptomatic medication (in our case, donepezil), but it seems that the task is difficult to achieve with the use of classical paper-pencil tests. The ability of computerized cognitive tests to detect a cognitive change due to a single dose of acetylcholinesterase inhibitors (AChEIs) in AD has not been extensively examined. Computerized (automated) cognitive tests have numerous advantages in comparison with classical “paper-pencil” neuropsychological tests and batteries. Computerized tests may provide more detailed results than the classical cognitive tests; they have multiple alternate test forms, and yield data that usually do not have floor or ceiling effects [10–13]. Furthermore, most tests in computerized batteries are based on visual stimuli and thus are language- and culture-independent. Some computerized tests, such as the Groton Maze Learning Test (GMLT), have been shown to be able to detect the change after a single dose of donepezil [12,14]. However, GMLT is a complex test that depends on several cognitive functions that are not the functions affected early by AD [15]. Very little is known about possibilities of other computerized tests more directly related to memory and learning to detect the effect of a single dose of donepezil. Therefore, we decided to investigate the potential ability

to detect significant cognitive change after a single dose of donepezil using a selection of Cambridge Neuropsychological Test Automated Battery (CANTAB) tests and analyzing a wide variety of different measures provided by these tests. CANTAB tests, especially the Paired Associates Learning (PAL) test, have many advantages in AD research, including early diagnosis of AD and amnesic MCI [16–19]. To the best of our knowledge, CANTAB tests have not been investigated for detection of the effect of a single dose of donepezil. In the future, with the increase of symptomatic treatment options, such psychometric tools for very early detection of treatment effect could be used in challenge tests for optimization of symptomatic treatment for an individual AD patient. With the advent of personalized medicine, the importance of individual selection of optimal treatment, tailored to a specific patient, becomes increasingly evident [20]. Personalization of treatment may be achieved by using various methods and strategies from genetic sequencing, neuroimaging studies, and sensitive electrophysiological methods, to direct and early evaluation of patient response to several treatment options [20–22]. The present study focused on the assessment of computerized cognitive tests and measures that can reliably detect change in cognition after a single dose of donepezil. In the future, similar cognitive instruments for personalization and optimization of AD treatment could be evaluated for other treatment options, thus leading to a selection of tools suitable to choose the optimal treatment for the individual AD patient.

We examined the change of cognitive functions in patients with AD in a prospective, rater-blinded, randomized study, using the CANTAB. The CANTAB is a validated, reliable neuropsychological battery [23], which consists of memory, learning, attention, problem solving, and executive function tests [24]. It measures associative memory and learning, with demonstrated specificity and sensitivity in detecting memory impairments in older adults [11,25]. We hypothesized that some measures of CANTAB tests would be able to detect the difference of change (greater improvement) of cognitive function in patients administered a single dose of donepezil after the first testing session compared to those who started their donepezil treatment after the second testing session and were still treatment-naïve between testing session 1 and testing session 2.

The idea to evaluate the effect of a single dose of medication is not new in medicine. This kind of attempt to elucidate the effects of a single dose of medication had various names, including challenge tests and pharmacological probes. Sometimes it was a single dose, and sometimes it was a short and intensive administration of the medication. The aim was to clarify the response of the disease under investigation to the specific medication, with the hope that the results of a single or short administration of medication could provide useful clues regarding efficacy of future long-term treatment. The concept

of challenge tests cannot be used in AD treatment at present due to the lack of available symptomatic treatment and very limited knowledge about suitable psychometric instruments to evaluate the results of challenge tests.

### Objectives

The objective the present study was to establish the ability of CANTAB tests and their measures to detect significant cognitive change after a single 5-mg dose of donepezil in treatment-naïve AD patients.

## Material and Methods

### Participants

This prospective, randomized, rater-blind study was performed at the Memory Disorders Unit of the Neurology Center, Vilnius University Hospital Santariskiu Clinics. We enrolled 92 subjects in the study. We recruited 62 consecutive, *de novo*-diagnosed, treatment-naïve AD patients and 30 healthy controls (Control group, CG) matched according to age, education, and gender. All patients were diagnosed with AD in usual clinical practice settings by a neurologist not involved in this study. Patients started their treatment with donepezil when the medication was prescribed by a neurologist of the Memory Disorders Unit. After the assessment day (the first day of dosing), the patients continued their treatment with donepezil as per usual clinical practice rules according to the treatment guidelines established by the Lithuanian Ministry of Health in the directives No. 382 and V-156. No modifications to the patient treatment were made due to this research.

### Study design

Informed consent was obtained, screening evaluation (including MMSE and GDS) was performed, inclusion/exclusion criteria were verified, and both sessions of CANTAB testing were performed on the same day when the patients took their first dose of donepezil. AD patients were randomly assigned to 1 of 2 research groups with the ratio 1: 1 using the sequence of random numbers 1 or 2, produced by the on-line Research Randomizer at <http://www.randomizer.org/>. Thirty AD patients were assigned to the AD+ group and 32 patients to the AD- group. Patients allocated to the AD+ group received a 5-mg donepezil tablet immediately after the first CANTAB testing session. The second testing session was performed 4 hours after the AD+ group patients took donepezil. The 4-hour period was selected because this corresponds to the time at which the peak plasma concentration of donepezil is observed after oral administration. Patients in the AD- group underwent both assessment sessions without taking donepezil between

sessions (i.e., both AD+ and AD- groups completed the first CANTAB testing session while treatment-naïve). The second CANTAB testing session was completed by the AD+ group 4 hours after the first single 5-mg dose of donepezil. The AD- group completed the second CANTAB testing session 4 hours after the first testing session, but were still treatment-naïve. The AD+ and AD- groups did not differ by age, education, or gender, as verified after the completion of the recruitment period. The neurologist performing CANTAB testing (JK) was blinded to the participant's assignment to the AD+, AD-, or control group. Randomization was accomplished; MMSE, GDS, and Hachinski ischemic score were assessed and donepezil usage instructions were provided by another neurologist (GK). Control group participants were recruited from the spouses, relatives, and accompanying persons of other patients attending the Neurology Department, but with no medical history of AD or other dementia.

### Approval by Ethics Committee

The study Protocol and Informed Consent Form (ICF) were approved by the Vilnius Regional Biomedical Research Ethics Committee (approval No. 158200-12-128-36). Written informed consent was obtained from all participants.

### Inclusion and exclusion criteria

Rigorous inclusion and exclusion criteria were applied for inclusion in the study. Inclusion and exclusion criteria for AD+ and AD- groups were the same. Randomization to AD+ or AD- groups was performed after enrollment.

Inclusion criteria for AD patients were:

- The patient has late-onset sporadic probable AD diagnosed according to the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer Disease and Related Disorders Association (NINCDS- ADRDA) criteria;
- The patient has mild or mild-to-moderate dementia: Mini-Mental State Examination (MMSE) score of at least 18, and no greater than 23;
- The patient is treatment-naïve (newly diagnosed AD);
- The patient has a CT or/and MRI in the process of AD diagnosis establishment with results consistent with the diagnosis of probable AD and without evidence of a major stroke;
- The patient is at least 65 years old;
- Hachinski ischemic score is equal to or less than 4;
- Geriatric Depression Scale (GDS) score is equal to or less than 19;
- Education is equal to or more than 8 years;
- The patient's sight and hearing are sufficient for compliance with the study assessment;
- The patient is proficient in the Lithuanian language.

Exclusion criteria for AD patients were:

- The patient is currently receiving or has taken any other cognition-enhancing medication within 6 months prior to the assessment;
- The patient has evidence of any neurodegenerative disease, or other serious neurological disorders other than AD, including, but not limited to, Lewy body dementia, fronto-temporal dementia, Parkinson disease, stroke, major head trauma, seizures, cerebral neoplasia, or systemic medical diseases that are likely to affect central nervous system functioning;
- The patient has tested positive for human immunodeficiency virus (HIV), hepatitis B surface antigen, or hepatitis C virus;
- The patient has a DSM-IV-TR Axis I disorder other than AD, including delirium, schizophrenia, schizoaffective disorder, bipolar disorder, current major depressive episode, or psychosis;
- The patient has evidence of clinically significant comorbidities, including, but not limited to, pulmonary, gastrointestinal, renal, hepatic, endocrine, cardiovascular system disease, or vitamin B<sub>12</sub> deficiency;
- Current or past alcohol or drug abuse.

Inclusion criteria for Control group participants were:

- Normal cognitive functioning (MMSE score 27–30);
- The patient is at least 65 years old;
- Hachinski ischemic score is equal to or less than 4;
- Geriatric Depression Scale (GDS) score is equal to or less than 19;
- Education is equal to or more than 8 years;
- The patient's sight and hearing are sufficient for compliance with the study assessment;
- The patient is proficient in the Lithuanian language.

Exclusion criteria for Control group participants were the same as for AD patients.

### Neuropsychological assessment instruments

Global cognitive performance of participants was assessed using the Lithuanian version of the Mini-Mental State Examination (MMSE). The Cambridge Neuropsychological Test Automated Battery (CANTAB®, Cambridge Cognition Ltd., UK) was used as the main instrument for detailed and more sensitive research assessment. CANTAB is a computer-based battery using a touch tone screen and press pad with 2 buttons. The order of test sequence remained constant across both test sessions, because 2 special batteries with the same sequence have been set-up of tests selected for this research according to the instructions provided by CANTAB Software User Guide. Separate batteries were assembled for testing 1 and testing 2; the only difference between them was that parallel versions of most tests (where available) were used at different testing sessions to minimize any potential learning effect.

After an initial explanation, subjects were given the following tests in the following order:

- Choice reaction time (CRT): a 2-stimuli visual discrimination and category achievement test;
- Stockings of Cambridge (SOC): The task is analogous to the 'Tower of London' test and assesses the subject's ability to engage in spatial problem solving. This test makes substantial demands on executive function;
- Paired associate learning (PAL): Assessment of simple visual pattern and visuospatial associative learning, which contains aspects of both a delayed response procedure and a conditional learning task;
- Pattern recognition memory (PRM) immediate (PRMi): A test of visual recognition memory in a 2-choice forced discrimination paradigm. The recognition task was performed immediately after a series of stimuli presentation;
- Spatial working memory (SWM): this task assesses the subject's ability to retain spatial information and to manipulate remembered items in working memory;
- Pattern recognition memory (PRM) delayed (PRMd): A test of visual recognition memory in a 2-choice forced discrimination paradigm. The recognition task was performed 30 minutes after stimuli presentation.

Raw scores of CANTAB test measures were selected for statistical assessment. Standard scores and the internal normative database of CANTAB were not used for comparison of results. Instead, our own control group was enrolled in the study because the internal normative database of CANTAB involved healthy volunteers and provides estimations with subjects matched for age, gender, and National Adult Reading Test (NART) scores. There is no analogue to NART in the Lithuanian language and it is hard or impossible to devise something similar due to the phonemic orthography used in the Lithuanian language. Therefore, the score of NART was left blank when entering initial participant data before testing. This invalidates the comparison with the internal normative database of CANTAB. Instead of NART, the education level in years was used as a proxy of premorbid intelligence level. It should be noted that while the learning effect was impossible due to use of parallel test versions in testing 1 and testing 2, the practice effect might have some influence on the results of testing 2.

### Statistical analysis

Comparisons between groups were performed using analysis of variance (ANOVA) for continuous variables and chi-square test for categorical variables. Normal distribution of data was verified using the Shapiro-Wilk test. One-way ANOVA with Bonferroni post-hoc test was used to assess differences between the results of the first testing session among the 3 participant groups. The Levene test was used to assess the homogeneity of variances across participant groups. The tables

provided below indicate when the assumption of homoscedasticity (homogeneity of variance) was violated.

Correlation of CANTAB test measures with MMSE scores was assessed using Pearson correlation coefficient  $r$ .

Repeated-measures ANOVA was applied to answer the question, whether the mean change in the cognitive function of the first to second testing session differed in the three groups. Scores on the CANTAB test measures at each testing session were submitted to a repeated-measures analysis of variance with test session number (first and second) entered as independent variables, and CANTAB test measure score entered as a dependent variable. The differences in change were measured directly by the "testing session" \* "group" interaction effect. The Bonferroni post hoc test was used for comparisons of the 3 independent groups. Sphericity of the variances of the differences between all possible pairs of groups was tested by using the Mauchly's sphericity test. The assumption of sphericity has not been violated, which could be due to the fact that there were only 2 levels of repeated measures. All 3 participant groups were included in the data set for repeated-measures ANOVA. Test results of 1<sup>st</sup> and 2<sup>nd</sup> testing session were included as 2 within-factor levels. Belonging to 1 of the 3 participant groups was entered as a between-factor.

The statistical significance value was set at  $p < 0.05$ .

## Results

### Demographic characteristics, depression level, and overall cognitive function

Study groups did not differ significantly by age ( $p=0.828$ ), education ( $p=0.952$ ), or gender ( $p=0.948$ ). Demographic characteristics, depression level by GDS, and MMSE scores for all groups are shown in Table 1.

### Comparison of CANTAB test measures in participant groups at Baseline (1<sup>st</sup> testing session)

Before proceeding to the main purpose of this study, it was necessary to establish which tests and which specific measures of the tests were able to distinguish AD patients from the Control group. For those tests that do not make this distinction, investigation of improvement is pointless. While these tests might be useful for research of cognitive enhancement, they are not relevant to AD research. One-way ANOVA was used to evaluate the significance of differences between participant groups. The Bonferroni post hoc test was used for multiple comparisons between separate groups. Normative data and standard scores (the number of standard deviations from

the mean) of the peer group provided by the CANTAB internal database were not used in our research due to a problem with NART scores explained above. Standard scores were invalidated by the absence of National Adult Reading Test (NART) data in our study. The control group, matched for age, education, and gender, was used for comparison purposes.

Results of CANTAB test measures for memory functions (PAL and PRM tests) are provided in Table 2.

Results of CANTAB test measures for other cognitive domains (CRT, SOC, and SWM tests) are provided in Table 3.

### Correlation between CANTAB test measures in first testing session and dementia severity (MMSE)

Correlations between CANTAB test measures and MMSE (as a measure of global dementia severity) were established to evaluate whether change in CANTAB test measures may be treated as a clinically relevant cognitive change (improvement). Only those test measures able to distinguish AD groups from the Control group were selected for correlation analysis. Eight measures of PAL and PRM tests showed statistically significant and strong correlation ( $r > 0.7$ ) with MMSE. The repeated measures ANOVA was performed with these test measures.

### Assessment of cognitive change due to a single dose of donepezil based on CANTAB test measures in participant groups (change between 1<sup>st</sup> and 2<sup>nd</sup> testing sessions)

For those CANTAB test measures that were able to distinguish AD and Controls and demonstrated the statistically significant and strong correlation ( $r > 0.7$ ) with MMSE (as a measure of global dementia severity, or global cognitive functioning), repeated-measures ANOVA was used to assess which measures are able to detect significant cognitive change due to a single dose of donepezil.

Significant interaction effect "Testing session" \* "Group" was established only for seven PAL test measures (Table 4). For all 7 PAL test measures Bonferroni post-hoc revealed that the difference is significant only between the results of the second testing session in AD+ and the second testing session in AD- group (between-group effect), while the first testing session results between AD+ and AD- groups were insignificant. If more demanding requirements for the ability of the test to detect the change are accepted, that is, significant change should be detected between AD+ group first testing and second testing results (within treatment group effect), then only 4 PAL test measures are able to fulfil this requirement. This means that only these 4 PAL test measures fulfil both criteria to be able to detect significant cognitive change after a single dose of donepezil: 1) significant difference of change in AD+

**Table 1.** Demographic characteristics, depression, and MMSE scores in participant groups.

	AD+ group	AD- group	Control group	Test
Number of subjects, N	30	32	30	
Age (years) Mean ±SD	77.30±5.11	77.03±5.28	76.43±6.36	ANOVA F=0.189; p=0.828 ns
Education (years) Mean ±SD	13.17±4.79	13.47±4.02	13.20±3.61	ANOVA F=0.050; p=0.952 ns
Gender Women/Men, N	17/13	17/15	17/13	Chi-square 0.106; p=0.948 ns
Depression (GDS score) Mean ±SD	7.67±4.93	6.84±3.91	6.77±4.34	ANOVA F=0.388; p=0.680 ns
MMSE score Mean ±SD	21.57±1.57	21.25±1.48	29.47±0.57	ANOVA F=393.5; p<0.001* Bonferroni post-hoc: AD+=AD-; CG>AD+; AD-

\* One-way ANOVA; ns – not significant.

**Table 2.** Comparison of Paired Associates Learning (PAL) and Pattern Recognition Memory (PRM) test measures in participant groups.

Measure	AD+ group	AD- group	Control group	One-way ANOVA	Bonferroni post-hoc
PAL First trial memory score	9.50±3.59	8.78±3.79	17.83±4.68	F= 47.08; p<0.001	CG>AD+,AD- AD+=AD-
PAL Mean errors to success	8.47±3.31	8.48±2.63	2.64±2.86	F=39.88; p<0.001	CG<AD+,AD- AD+=AD-
PAL Mean trials to success	3.67±0.68	3.74±0.76	1.85±0.71	F=68.14; p<0.001	CG<AD+,AD- AD+=AD-
PAL Stages completed	6.30±1.12	6.00±0.88	7.93±0.25	F=47.27; p<0.001	CG>AD+,AD- AD+=AD-
PAL Stages completed on first trial	3.60±1.28	3.50±1.19	5.40±1.22	F=23.01; p<0.001	CG>AD+,AD- AD+=AD-
PAL Total errors (adjusted)	107.3±36.6	119.2±32.2	20.83±22.3	F=91.42; p<0.001	CG<AD+,AD- AD+=AD-
PAL Total errors (1 shape, adjusted)	0.10±0.40	0.34±0.83	0.00±0.00	F=3.349*; p<0.05	Comparison not valid
PAL Total errors (2 shapes, adjusted)	2.10±2.39	2.15±3.07	0.33±1.02	F=5.930*; p<0.05	Comparison not valid
PAL Total errors (3 shapes, adjusted)	12.63±12.5	13.15±12.2	1.20±1.39	F=13.43*; p<0.001	Comparison not valid
PAL Total errors (6 shapes, adjusted)	33.00±15.2	38.71±13.8	6.07±6.87	F=59.09; p<0.001	CG<AD+,AD- AD+=AD-
PAL Total errors (8 shapes, adjusted)	59.50±16.9	64.84±14.5	13.23±14.1	F=105.9; p<0.001	CG<AD+,AD- AD+=AD-
PAL Total trials (adjusted)	31.67±7.16	32.75±7.27	14.57±5.08	F=72.59; p<0.001	CG<AD+,AD- AD+=AD-
PRM immediate; Number correct	8.03±1.47	7.44±1.65	9.93±1.39	F=22.84; p<0.001	CG>AD+,AD- AD+=AD-
PRM delayed; Number correct	6.30±1.84	6.43±1.39	9.40±1.69	F=34.25; p<0.001	CG>AD+,AD- AD+=AD-

Measures of CANTAB tests provided as Mean ±SD of the Raw Score; \* Levene test significant (P<0.05), ANOVA invalidated.

**Table 3.** Comparison of Choice Reaction Time (CRT), Stockings of Cambridge (SOC), and Spatial Working Memory (SWM) test measures in participant groups.

Measure	AD+ group	AD- group	Control group	One-way ANOVA	Bonferroni post-hoc
CRT Mean correct latency (ms)	573.0±192.9	548.0±166.8	536.6±230.9	F=0.266; p=0.767 ns	CG=AD+=AD-
CRT Total correct trials (N)	98.97±1.40	98.72±2.14	98.50±2.01	F=0.459; p=0.633 ns	CG=AD+=AD-
CRT Total incorrect trials (N)	0.87±1.20	1.00±1.81	1.20±1.64	F=0.338; p=0.714 ns	CG=AD+=AD-
SOC Mean moves (2 moves minimum)	2.15±0.35	2.28±0.51	2.03±1.83	F=3.394*; p<0.05	Comparison not valid
SOC Mean moves (3 moves minimum)	3.75±0.73	3.55±0.65	3.16±0.37	F=7.164; p=0.001	CG<AD+, AD-AD+=AD-
SOC Mean moves (4 moves minimum)	5.11±0.98	5.15±1.04	4.90±0.93	F=0.576; p=0.564 ns	CG=AD+=AD-
SOC Mean moves (5 moves minimum)	6.60±1.65	7.03±1.38	6.61±1.51	F=0.820; p=0.444 ns	CG=AD+=AD-
SWM Total errors	60.3±14.5	64.1±10.5	36.2±17.2	F=34.29; p<0.001	CG<AD+, AD-AD+=AD-
SWM Total errors (4 boxes)	5.03±3.58	6.12±3.35	2.37±2.31	F=11.66; p<0.001	CG<AD+, AD-AD+=AD-
SWM Total errors (6 boxes)	20.0±6.53	19.7±6.58	11.0±7.89	F=16.17; p<0.001	CG<AD+, AD-AD+=AD-
SWM Total errors (8 boxes)	35.2±7.86	38.2±6.67	22.8±9.01	F=32.78; p<0.001	CG<AD+, AD-AD+=AD-

Measures of CANTAB tests provided as Mean ±SD of the Raw Score; \* Levene test significant (P<0.05), ANOVA invalidated; ns – not significant.

and AD- groups; and 2) significant change of mean in absolute values between testing 1 and testing 2 in AD+ group due to a single dose of donepezil (within treatment group effect). These 4 best PAL test measures are: PAL Mean trials to success, Total errors (adjusted), Total errors (6 shapes, adjusted), Total trials (adjusted).

Results of interaction effect (“Testing session” \* “Group”), between group effect on second testing, and within AD+ group effect are provided in Table 4. It should be noted that Table 4 contains only indicated results, but not all effects assessed by repeated measures ANOVA.

Only some comparisons provided by repeated-measures ANOVA are presented in Table 4. The patterns of change between testing 1 and testing 2 are very variable for different test measures. Figure 1A shows the changes in PAL First trial memory score results. This measure meets criterion 1, but fails to meet criterion 2. Figure 1B and 1C illustrate the best PAL test measures (2 of 4) – PAL Mean trials to success and Total errors (adjusted). Both PAL test measures fulfil both criteria. Figure 1D illustrates that the change of PRM test measure “PRM delayed

(Number correct)” was very different. This PRM test measure meets the criterion 2, but fails to meet criterion 1. This variety of changes in CANTAB test results is more explicitly discussed in the Discussion section.

## Discussion

The Choice Reaction Time (CRT) test measures speed of response in a simple 2-choice paradigm using a 2-button press pad [26]. Our results did not show a difference in CRT test performance in mild and mild-to-moderate AD patients and the Control group. The CRT test measures speed of response in a simple 2-choice paradigm. It seems that the test is too simple and easy for all participant groups. A clear ceiling effect is visible in the results of this test, but this does not necessarily mean that the speed of psychomotor reactions is not altered in AD patients. Even the CRT may be informative in moderate and severe AD patients, but for our participants, CRT is unable to distinguish AD patients from normal controls. These results could not be extrapolated to other mild dementias, especially those related to dementia plus Parkinsonism syndromes, such

**Table 4.** Results of repeated-measures ANOVA with Bonferroni post-hoc in participant groups based on first and second testing sessions\*.

Measure	Difference of changes in AD+ and AD- groups between 1 <sup>st</sup> and 2 <sup>nd</sup> testing (interaction effect)	p	Difference of AD+ and AD- groups based on the results of the second testing (between group effect)	p	Difference of first and second testing results in AD+ group (within group effect)	p
PAL First trial memory score	3.39	0.0016	4.11	<0.001	1.80	0.130 ns
PAL Mean trials to success	-0.59	0.0014	-0.62	0.003	-0.51	<0.001
PAL Stages completed	0.95	<0.0001	1.25	<0.001	0.23	1.00 ns
PAL Total errors (adjusted)	-35.4	<0.0001	-47.3	<0.001	-18.3	0.005
PAL Total errors (6 shapes, adjusted)	-11.7	0.0002	-17.5	<0.001	-7.03	0.008
PAL Total errors (8 shapes, adjusted)	-8.89	0.0447	-14.2	0.005	-4.67	1.00 ns
PAL Total trials (adjusted)	-8.65	<0.0001	-9.73	<0.001	-4.40	0.002
PRM delayed; Number correct	0.61	0.240 ns	0.47	1.00 ns	1.20	0.003

\* Only those CANTAB test measures were included, which correlated strongly ( $r>0.7$ ) with MMSE; ns – not significant.

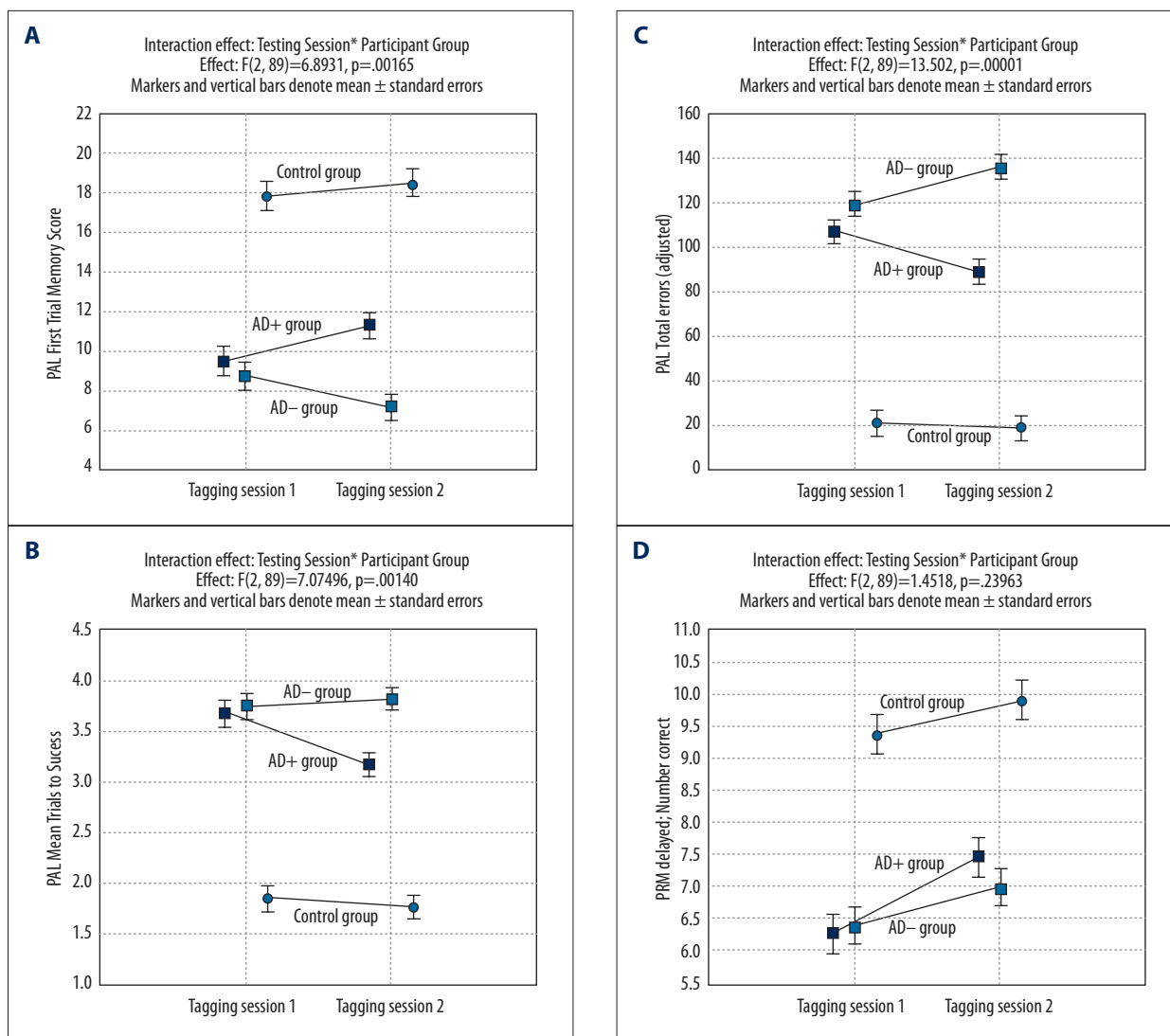
as dementia with Lewy bodies (DLB) and Parkinson disease dementia (PDD). However, the CRT is not suitable for evaluation of attention or psychomotor speed in mild AD. The positive implications that can be drawn from our results with the CRT test are that other cognitive results of our study could not be attributed to malfunction of attention systems or psychomotor speed. Deterioration in other cognitive functions, detected by other tests, is independent of attention in our study, as the results of the test for attention (CRT) were normal in AD groups. This allows more confidence in drawing conclusions about other test results, because disorders of attention seriously distort the results of any other cognitive tests. As no clear disorder of attention was detected in our study, memory, executive, and other cognitive disorders can confidently be attributed to the corresponding cognitive systems.

The Stockings of Cambridge (SOC) test assesses spatial planning and motor control. This test gives a measure of frontal lobe function [26]. The results of SOC test in our study are interesting, unequivocal, and clearly need further investigation. The SOC test is used for evaluation of complex executive functions, working memory, and planning, which are attributed to the frontal lobe. Published reports about frontal function in AD are quite controversial. Some of them indicate early abnormalities of executive functions in AD [27,28], and others did not show a significant difference between AD patients and normal controls [11]. Problems that can be solved in 2 minimum

moves are easy for AD patients and controls and provide inconclusive results. Problems that can be solved in 4 or 5 minimum moves are quite difficult for AD patients and controls. The “informative window” is very narrow – only problems that could be solved in 3 minimum moves clearly distinguished AD patients and controls (Table 3). Although this result is interesting in itself, it shows how easy is to pass over indicators of frontal dysfunction in AD. Moreover, correlation of the results of the “3 moves task” with overall dementia severity by MMSE is quite weak, albeit significant. We performed repeated-measures ANOVA for this task of the SOC test (results are not provided in this article), but the difference of changes in AD+ and AD- was not significant. These results show that frontal executive functions are affected in mild AD, but research in this field is very demanding and studies should be very cautious and well designed. Importantly, the effect of cholinergic treatment on executive dysfunction in AD is much smaller than the effect on memory, which raises a number of interesting questions, such as “Is executive dysfunction in AD dependent on cholinergic deficit? If not, what is the neurochemical basis of this dysfunction?”. Summarizing, it should be stressed that the investigation of frontal dysfunction in AD is far from over and extensive research is needed to elucidate the place and mechanisms of executive (frontal) dysfunction in AD.

The Spatial Working Memory (SWM) test assesses working memory and strategy use. This test is a sensitive measure of





**Figure 1.** Changes between testing sessions 1 and 2 in participant groups of PAL First Trial Memory Score, PAL Mean Trials to Success, PAL Total errors (adjusted), and PRM delayed number correct score (A–D).

working memory, frontal lobe, and executive dysfunction [26]. While some, but not all, SWM test measures are significantly worse in AD than in normal controls (Table 3, only some significant SWM measures are provided), SWM test results showed only a moderate correlation with the MMSE. Repeated-measures ANOVA with the SWM test provided inconsistent and patchy results. Many of our comments about the SOC test also apply to the SWM, but while the SOC test depends heavily on planning abilities, the SWM places heavy demands on working memory. Comparison of change in AD+ and AD- groups after donepezil administration indicates that SWM is not significantly dependent on the cholinergic status of the brain. Possible use of the SWM test in AD research needs further investigation.

The Pattern Recognition Memory (PRM) test assesses visual recognition memory. PRM is a test of visual pattern recognition

memory in a 2-choice forced discrimination paradigm. This test is sensitive to dysfunction in medial temporal areas of the brain and is relatively insensitive to dysfunction in the frontal lobe. PRM is a relatively simple test. Although it is a memory test, as a test for assessment of recognition memory it has different significance in AD than recall memory tests. Recognition memory is relatively well preserved in AD, while disorders of episodic recall memory are a hallmark of AD. Our results showed that the PRM test provided significantly worse results in both AD groups than in the control group (Table 2) and had a strong correlation with MMSE, but repeated-measures ANOVA did not show significant difference of change in AD+ and AD- groups after donepezil challenge (Table 4, Figure 1D), even though the change in AD+ score from testing session 1 to testing session 2 was significant. Because the dynamics of performance of the Control group on the PRM test is

similar to AD-, a conclusion could not be drawn about memory change itself, rather it might be hypothesized that a single dose of donepezil somehow increased retaining of attention or decreased the tiredness in AD+ group.

The Paired Associates Learning (PAL) test assesses episodic visual recall memory and new learning [26]. This test is primarily sensitive to changes in medial temporal lobe functioning. PAL is by far the best test in mild and mild-to-moderate AD. These results are in line with previously published findings [29–32]. Most of the PAL test measures provided significantly worse results in both AD groups than in the Control group (Table 2) and had a strong correlation with MMSE. Only measures of the PAL test showed a significant difference of change in AD+ and AD- groups after donepezil challenge (Figure 1A–1C). Some measures of the PAL test were able to detect significant differences of change in AD+ and AD- and statistically significant improvement in the results of AD+ group itself due to a single dose of donepezil (Table 4), while this improvement was absent in the AD- group (Figure 1A–1C). This indicates that such PAL test measures as PAL Mean trials to success, PAL Total errors (adjusted), PAL Total errors (6 shapes, adjusted), and PAL Total trials (adjusted) are reliable for measurements of the effect in donepezil challenge tests. PAL First trial memory score, PAL Stages completed, and PAL Total errors (8 shapes, adjusted) showed significant differences of change in AD+ and AD- groups, but failed to demonstrate significant changes in the AD+ group itself. Again, the situation is ambiguous – performance due to donepezil did not increase significantly (AD+ group), but decrease of performance without donepezil (AD- group) contributed substantially to the overall result of significant difference in change. Tiredness or different ability to retain attention may be responsible for this result. Either way, these CANTAB test measures are not the best to demonstrate the clinically relevant real change in memory itself. We deliberately provide 4 separate graphs in Figure 1 – not for all 4 of the best tests, because their response patterns are quite similar (only 2 of them are shown in Figure 1 – PAL Mean trials to success and PAL Total errors (adjusted), but the graphs of change between testing session 1 and testing session 2 – which demonstrate the different patterns of change discussed above.

## References:

1. Sosa-Ortiz AL, Acosta-Castillo I, Prince MJ: Epidemiology of dementias and Alzheimer's disease. *Arch Med Res*, 2012; 43: 600–8
2. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM: Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement*, 2007; 3: 186–91
3. Whitehouse PJ, Price DL, Struble RG et al: Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science*, 1982; 215: 1237–39
4. Whitehouse PJ, Price DL, Clark AW et al: Alzheimer disease: evidence for selective loss of cholinergic neurons in the nucleus basalis. *Ann Neurol*, 1981; 10: 122–26
5. Hasselmo ME: The role of acetylcholine in learning and memory. *Curr Opin Neurobiol*, 2006; 16: 710–15
6. Rogers SL, Doody RS, Mohs RC, Friedhoff LT: Donepezil improves cognition and global function in Alzheimer disease: a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. *Arch Intern Med*, 1998; 158: 1021–31
7. Rogers SL, Farlow MR, Doody RS et al: A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology*, 1999; 50: 136–45
8. Bryson HM, Benfield P: Donepezil. *Drugs Aging*, 1997; 10(3): 234–39

It should be noted that our study leaves many questions unanswered. It is not clear whether the initial impressive improvement in PAL test results after donepezil challenge will be durable in long-term treatment and after down-regulation of acetylcholine receptors. It is not clear whether improvement of PAL results after a single dose of donepezil will correlate significantly with overall cognitive function based on MMSE during long-term treatment. The question remains open whether the initial donepezil challenge effect on PAL test results can distinguish responders to donepezil from non-responders after a longer period of treatment. Many remaining questions require follow-up study with the same study population. Moreover, AD in real clinical practice usually is accompanied by cerebrovascular disease and diabetes, which may contribute to the pathogenesis of the AD itself, and also can change response to cholinergic treatment [33]. Our results are important because they indicate there are at least several CANTAB test measures able to detect significant change after a single dose of donepezil. Having more options to evaluate the results of pharmacological challenge in cognitive neurology may be useful in future clinical trials of new symptomatic treatment for AD, may help to identify responders early in the course of treatment, and could provide reliable tools for personalization of AD treatment.

## Conclusions

Four CANTAB PAL test measures are able to detect reliable and clinically relevant cognitive change after a single 5-mg dose of donepezil in AD patients: PAL Mean trials to success, PAL Total errors (adjusted), PAL Total errors (6 shapes, adjusted), and PAL Total trials (adjusted). These 4 CANTAB PAL test measures detected a significant difference in cognitive performance change in the AD+ group in comparison with the AD- and Control groups. Mean scores within the AD+ group itself improved significantly following a single 5-mg dose administration of donepezil.

## Conflict of interest

The authors have no conflicts of interest to declare.

9. Tariot PN, Cummings JL, Katz IR et al: A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting. *J Am Geriatr Soc*, 2001; 49: 1590-99
10. Collie A, Darekar A, Weissgerber G et al: Cognitive testing in early-phase clinical trials: development of a rapid computerized test battery and application in a simulated Phase I study. *Contemp Clin Trials*, 2007; 28: 391-400
11. Egerházi A, Berecz R, Bartók E, Degrell I: Automated Neuropsychological Test Battery (CANTAB) in mild cognitive impairment and in Alzheimer's disease. *Prog Neuropsychopharmacol Biol Psychiatry*, 2007; 31(3): 746-51
12. Pietrzak RH, Maruff P, Snyder PJ: Methodological improvements in quantifying cognitive change in clinical trials: an example with single-dose administration of donepezil. *J Nutr Health Aging*, 2009; 13(3): 268-73
13. Yurko-Mauro K, McCarthy D, Rom D et al: Beneficial effects of docosahexaenoic acid on cognition in age-related cognitive decline. *Alzheimers Dement*, 2010; 6(6): 456-64
14. Snyder PJ, Bednar MM, Cromer JR, Maruff P: Reversal of scopolamine-induced deficits with a single dose of donepezil, an acetylcholinesterase inhibitor. *Alzheimers Dement*, 2005; 1: 126-35
15. Pietrzak RH, Maruff P, Mayes LC et al: An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Arch Clin Neuropsychol*, 2008; 23: 433-45
16. Fowler KS, Saling MM, Conway EL et al: Computerized neuropsychological test in the early detection of dementia: prospective findings. *J Int Neuropsychol Soc*, 1997; 3: 139-46
17. Fowler KS, Saling MM, Conway EL et al: Paired associate performance in the early detection of DAT. *J Int Neuropsychol Soc*, 2002; 8: 58-71
18. Gould RL, Brown RG, Owen AM et al: Functional neuroanatomy of successful Paired Associate Learning in Alzheimer's disease. *Am J Psychiatry*, 2005; 162: 2049-60
19. Jakala P, Sirvio J, Riekkinen M et al: Guanfacine and clonidine, alpha 2-agonists, improve paired associates learning, but not delayed matching to sample, in humans. *Neuropsychopharmacology*, 1999; 20: 119-30
20. Stefano GB, Kream RM: Personalized- and one- medicine: bioinformatics foundation in health and its economic feasibility. *Med Sci Monit*, 2015; 21: 201-4
21. Zimny A, Bladowska J, Neska M et al: Quantitative MR evaluation of atrophy, as well as perfusion and diffusion alterations within hippocampi in patients with Alzheimer's disease and mild cognitive impairment. *Med Sci Monit*, 2013; 19: 86-94
22. Vaitkevičius A, Kaubrys G, Audronytė E: Distinctive effect of donepezil treatment on P300 and N200 subcomponents of auditory event-related evoked potentials in Alzheimer disease patients. *Med Sci Monit*, 2015; 21: 1920-27
23. Wild K, Howieson D, Webbe F et al: Status of computerized cognitive testing in aging: a systematic review. *Alzheimer's Dement*, 2008; 4: 428-37
24. Robbins TW, James M, Owen AM et al: Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia*, 1994; 5: 266-81
25. de Jager CA, Milwain E, Budge M: Early detection of isolated memory deficits in the elderly: the need for more sensitive neuropsychological tests. *Psychol Med*, 2002; 32: 483-91
26. CANTABclipse Test Administration Guide. Manual version 3.0.0. Cambridge Cognition Limited, 2006
27. Arnaiz E, Almkvist O: Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. *Acta Neurol Scand*, 2003; 179: 34-41
28. Twamley EW, Ropacki SA, Bondi MW: Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *J Int Neuropsychol Soc*, 2006; 12: 707-35
29. Juncilla J, Oja S, Laine M, Karrasch M: Applicability of the CANTAB-PAL computerized memory test in identifying amnesic mild cognitive impairment and Alzheimer's disease. *Dement Geriatr Cogn Disord*, 2012; 34(2): 83-89
30. Pike KE, Rowe CC, Moss SA, Savage G: Memory profiling with paired associate learning in Alzheimer's disease, mild cognitive impairment, and healthy aging. *Neuropsychology*, 2008; 22(6): 718-28
31. Ahmed S, Mitchell J, Arnold R et al: Predicting rapid clinical progression in amnesic mild cognitive impairment. *Dement Geriatr Cogn Disord*, 2008; 25(2): 170-77
32. Blackwell AD, Sahakian BJ, Vesey R et al: Detecting dementia: novel neuropsychological markers of preclinical Alzheimer's disease. *Dement Geriatr Cogn Disord*, 2004; 17(1-2): 42-48
33. Wang F, Guo X, Shen X et al: Vascular dysfunction associated with type 2 diabetes and Alzheimer's disease: a potential etiological linkage. *Med Sci Monit Basic Res*, 2014; 20: 118-29