

Featured Article

Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: A secondary analysis of the ADCS vitamin E and donepezil in MCI study

Emily C. Edmonds^{a,b,*}, M. Colin Ard^c, Steven D. Edland^{c,d}, Douglas R. Galasko^{a,b,c},
David P. Salmon^c, Mark W. Bondi^{a,b}

^aVeterans Affairs San Diego Healthcare System, San Diego, CA

^bDepartment of Psychiatry, University of California San Diego, School of Medicine, La Jolla, CA

^cDepartment of Neurosciences, University of California San Diego, School of Medicine, La Jolla, CA

^dDepartment of Biostatistics & Department of Family and Preventative Medicine, University of California San Diego, La Jolla, CA

Abstract

Introduction: Criteria for mild cognitive impairment (MCI) used in many clinical trials are susceptible to “false-positive (FP)” errors that can be avoided by an actuarial psychometric approach.

Methods: Cluster analysis was applied to baseline neuropsychological test data from 756 MCI participants in the Alzheimer's Disease Cooperative Study donepezil trial. Treatment groups were compared after FP MCI cases were removed.

Results: Cluster analyses revealed three groups: “single-domain amnesic MCI” (31%), “multi-domain amnesic MCI” (39%), and “FP MCI” (30%). After removing FP MCI cases, the donepezil treatment group had a lower rate of progression to Alzheimer's disease and better performance on cognitive tests than the placebo/vitamin E group.

Discussion: Removal of FP MCI diagnoses unmasked beneficial effects of donepezil, despite a 30% reduction in sample size. MCI subject selection based on actuarial methods with comprehensive neuropsychological test data can result in more efficient clinical trials and improved ability to detect treatment effects.

Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords:

Mild cognitive impairment; MCI; Alzheimer's disease; Dementia; Donepezil; Treatment; Neuropsychology; Misdiagnosis; False positive; Cluster analysis

1. Introduction

Mild cognitive impairment (MCI) is a transitional step between normal cognition and dementia in those with Alzheimer's pathology and is therefore a stage of Alzheimer's disease (AD) where interventions may prove useful for preventing or delaying progression to dementia [1,2]. The conventional criteria for MCI, when implemented for multicenter studies such as clinical trials targeting MCI [3]

and large-scale observational studies such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) [4], have included: subjective cognitive complaints, an impaired score on a single objective memory test, clinical judgment of cognitive decline but not dementia, and intact functional abilities [5,6]. Recent research suggests, however, that this diagnostic approach to MCI may be overinclusive. When we examined performance across a battery of cognitive tests by those with conventionally diagnosed MCI using an actuarial psychometric approach with normative data and cluster analysis techniques, we found that a large subgroup (e.g., >30% of the MCI cohort in ADNI) performed within normal limits, suggesting they may represent “false-positive (FP)” diagnostic errors [7,8]. This impression was

The authors have declared that no conflict of interest exists.

*Corresponding author. Tel.: +1 858-552-8585, ext. 6369; Fax: +1 858-626-8537.

E-mail address: ecedmonds@ucsd.edu or mbondi@ucsd.edu

<https://doi.org/10.1016/j.trci.2017.11.001>

2352-8737/Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

strengthened when we found that the FP MCI subgroup in the ADNI cohort had normal CSF amyloid β and tau biomarkers [8], normal positron emission tomography amyloid burden [9], normal cortical thickness measures [10], a low rate of progression to AD [8], and a high rate of reversion to “cognitively normal” within a few years [8].

The susceptibility of the conventional diagnostic approach to false positive MCI classification has major implications for clinical trials that target this population. The inadvertent inclusion of substantial numbers of cognitively normal, “disease-free” individuals in MCI cohorts involved in such trials could greatly weaken or mask meaningful results. We examined this possibility in the present study by reexamining the results of the Alzheimer’s Disease Cooperative Study (ADCS) vitamin E and donepezil trial in MCI [3], after using our actuarial psychometric approach to identify and remove subjects who were classified as FP MCI. The ADCS donepezil trial, conducted between March 1999 and January 2004, was a 36-month, multicenter, randomized, double-blind, placebo-controlled study of the effects of donepezil, vitamin E, or placebo on cognitive and functional decline in participants with amnesic MCI [3]. The original results showed no difference between groups in the rate of progression to AD after 36 months, although progression to AD was lower in the donepezil group relative to the placebo and vitamin E groups during the first 12 months of treatment, and that effect persisted for 24 months in the apolipoprotein E (*APOE*) $\epsilon 4$ carrier group [3]. We hypothesized that potential beneficial effects of donepezil were attenuated by the inclusion of subjects with a FP MCI diagnosis in the trial and that identification and removal of these subjects would strengthen the observed effect of donepezil on cognitive performance and progression to AD.

2. Methods

2.1. Participants and procedure

Details of subject selection, randomization, clinical evaluation, neuropsychological assessment, and other trial procedures have been published [3]. The ADCS donepezil trial randomized 769 participants who met the following diagnostic criteria for MCI: (1) a memory complaint corroborated by an informant, (2) abnormal memory function defined as scoring below the education-adjusted normative cutoff value on one paragraph from the Wechsler Memory Scale–Revised Logical Memory II subtest, (3) a Mini–Mental State Examination (MMSE) score of 24–30, (4) a global Clinical Dementia Rating score of 0.5, and (5) general cognition and functional ability sufficiently preserved so that a diagnosis of AD or dementia could not be made [3]. All subjects were clinically evaluated and underwent neuropsychological testing at baseline and every 6 months thereafter. The primary outcome was development of probable or possible AD according to NINCDS-ADRDA criteria, and neuropsychological performance was also assessed. Neuropsychological

data from the baseline assessment were incomplete for 13 participants. Therefore, our current analyses were based on 756 participants. The original study was approved by the relevant institutional review boards, and written informed consent was obtained from all participants. Data used for the current report were reanalyzed with permission.

2.2. Statistical analyses

Following our previous methods [8], six baseline neuropsychological test scores were converted into age-adjusted z-scores based on regression coefficients derived from a group of healthy control participants ($n = 112$) and entered into a hierarchical cluster analysis using Ward’s method. The six test scores included two measures of attention/executive function (Symbol-Digit Modalities and Backward Digit Span), two language measures (Boston Naming Test and Category Fluency), and two measures of memory (Alzheimer’s Disease Assessment Scale–Cognitive subscale [ADAS-Cog] Immediate and Delayed Word Recall). Resulting cluster-derived groups were compared using analyses of variance (ANOVAs) and chi-square tests with Bonferroni corrected post hoc comparisons (three cluster-derived group comparisons; $P = .05/3 = .02$).

Replicating the analysis conducted in the original study [3], a Cox proportional hazards model controlling for baseline variables (age, MMSE score, and *APOE* genotype) was used to examine time to the development of AD. Hazard ratios compared risk of progression to AD in the donepezil versus placebo/vitamin E group (the placebo and vitamin E groups were collapsed because there were no significant differences between them on any measure examined). A chi-square test was used to compare overall rates of progression to AD over the course of the 3 years.

ANOVAs were used to compare the donepezil group to the placebo/vitamin E group on neuropsychological measures used in the original study [3] at the 12-, 24-, and 36-month time points. No correction for multiple comparisons was applied, consistent with the original analysis [3]. ANOVAs were also used to examine interactions between group and *APOE* $\epsilon 4$ status on cognitive measures. All analyses were performed twice: (1) with all MCI participants (i.e., “original MCI sample”; $n = 756$) and (2) with the MCI sample that remained after FP MCI participants were excluded (i.e., “new MCI sample”; $n = 530$).

3. Results

3.1. Characteristics of cluster-derived MCI groups

Cluster analysis identified three subgroups (Fig. 1). A “single-domain amnesic MCI” group (aMCI-sd; $n = 235$; 31%) performed in the impaired range (>1.5 standard deviations [SDs] below mean) only on memory measures. A “multi-domain amnesic MCI” group (aMCI-md; $n = 295$; 39%) performed in the impaired range (>1.5 SDs below mean) on memory measures and had several mildly impaired scores (>1 SDs below mean) on measures of executive function

and language. A FP MCI group ($n = 226$; 30%) scored in the average to low-average range on all six neuropsychological measures despite their original diagnosis of MCI (note that none of the six measures were used in making the MCI diagnosis).

Demographic, clinical, and neuropsychological characteristics of the cluster groups are shown in Table 1. The FP MCI group was significantly younger and more educated than the other two groups. The FP MCI group had a lower rate of progression to dementia (5.8%) over the course of the trial in comparison to the other two MCI subgroups (35%–38%). The FP MCI group had fewer individuals who carried the *APOE* $\epsilon 4$ allele and greater independence in activities of daily living relative to the other MCI subgroups. In addition, they had significantly better scores on diagnostic and neuropsychological measures. The three cluster groups had comparable proportions of participants in the donepezil, vitamin E, and placebo arms of the trial.

3.2. Comparison of donepezil and placebo/vitamin E groups

In both the *original MCI sample* and the *new MCI sample*, there were no differences in age, education, or gender between the donepezil and placebo/vitamin E groups, and the groups had equivalent performance at baseline on all neuropsychological measures.

3.2.1. Original MCI sample

Consistent with the original report [3], 27.5% of participants (208/756) progressed to possible or probable AD by

month 36. There was no significant difference in the rate of development of dementia over 36 months between the donepezil group (59/247; 23.9%) and the placebo/vitamin E group (149/509; 29.3%), $\chi^2(1) = 2.42, P = .12$. A Cox proportional hazards analysis showed no significant difference in the probability of progression from MCI to AD between the donepezil and the placebo/vitamin E groups (hazard ratio = 0.78; 95% confidence interval = 0.58–1.05; $P = .10$).

Examination of neuropsychological test data showed that the donepezil group had higher ADAS-Cog total scores relative to the placebo/vitamin E group at the 12-month ($P = .02, d = 0.22$) visit; no differences were seen at 24 or 36 months. This difference at 12 months was largely driven by better performance of the donepezil group on immediate recall of a 10-item word list (ADAS-Cog Immediate Word Recall; see Table 2 and Fig. 2A). On an additional measure of immediate memory (New York University [NYU] Paragraph Immediate Recall), the donepezil and placebo/vitamin E groups did not differ at any time point (Table 2 and Fig. 2B). There were no significant group by *APOE* $\epsilon 4$ status interactions on ADAS-Cog total score or immediate memory performance. There were also no group differences on delayed recall of the words (ADAS-Cog Delayed Word Recall) or the paragraph (NYU Paragraph Delayed Recall).

3.2.2. New MCI sample (FP MCI removed)

In the new sample, the overall rate of progression to possible or probable AD over 36 months was 36.8% (195/530). The rate of progression was significantly lower in the donepezil group (52/174; 29.9%) than the placebo/vitamin

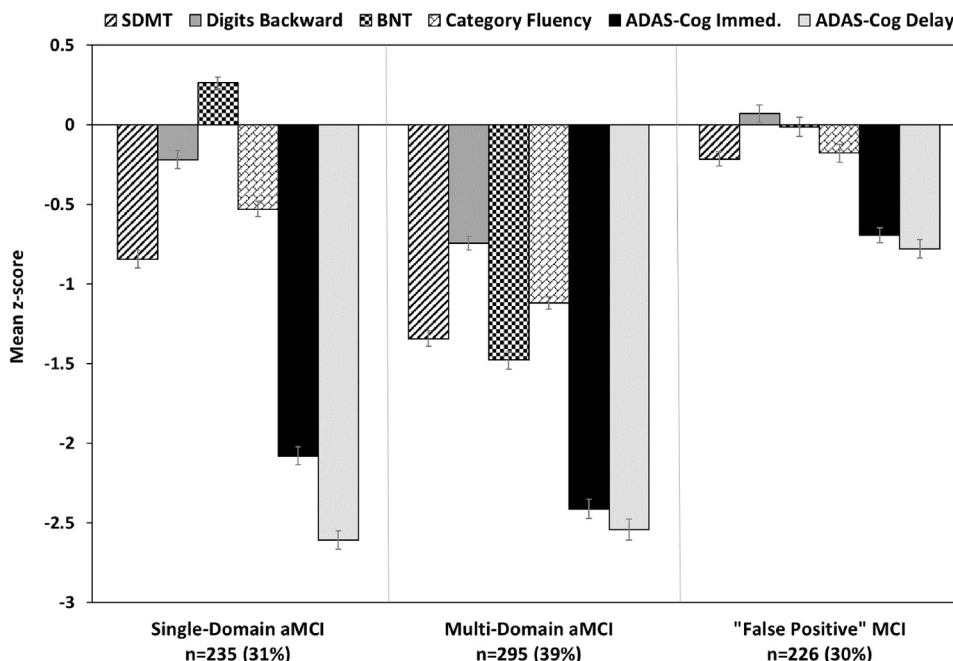


Fig. 1. Neuropsychological performance for the cluster groups. Error bars denote standard error of the mean. Abbreviations: aMCI, amnesic mild cognitive impairment; ADAS-Cog, Alzheimer's Disease Assessment Scale-Cognitive Subscale; BNT, Boston Naming Test; SDMT, Symbol-Digit Modalities Test.

Table 1
Demographic, clinical, and neuropsychological characteristics of the cluster groups

	Single-domain aMCI (n = 235)	Multi-domain aMCI (n = 295)	"False-positive" MCI (n = 226)	<i>F</i> or χ^2	<i>P</i> value	Effect size
Demographics						
Age (years)	73.5 (6.6)	73.7 (7.3)	71.7 (7.7)	<i>F</i> = 5.88	.003 ^{†,‡}	η_p^2 = .02
Education (years)	14.9 (2.8)	13.6 (3.4)	15.7 (2.6)	<i>F</i> = 33.24	<.001 ^{*,†,‡}	η_p^2 = .08
Gender (% male)	60.4%	49.8%	54.0%	χ^2 = 5.94	.05*	ϕ_c = .09
Clinical variables						
% progression to AD	35.3%	38.0%	5.8%	χ^2 = 77.00	<.001 ^{†,‡}	ϕ_c = .32
ADL Scale	45.7 (4.3)	45.1 (5.4)	47.4 (4.0)	<i>F</i> = 15.37	<.001 ^{†,‡}	η_p^2 = .04
% <i>APOE</i> ϵ 4 carriers	63.8%	58.0%	42.0%	χ^2 = 23.80	<.001 ^{†,‡}	ϕ_c = .18
Diagnostic measures						
LM II Recall	2.9 (2.5)	2.5 (2.1)	4.8 (2.2)	<i>F</i> = 62.62	<.001 ^{†,‡}	η_p^2 = .16
MMSE	27.2 (1.8)	26.5 (1.7)	28.3 (1.5)	<i>F</i> = 72.69	<.001 ^{*,†,‡}	η_p^2 = .16
CDR Sum of Boxes	1.9 (0.8)	2.0 (0.8)	1.5 (0.7)	<i>F</i> = 23.70	<.001 ^{†,‡}	η_p^2 = .06
Neuropsychological scores included in cluster analysis						
Symbol-Digit Modalities	31.3 (9.9)	25.8 (9.3)	39.2 (8.4)	<i>F</i> = 134.88	<.001 ^{*,†,‡}	η_p^2 = .26
Backward Digit Span	6.5 (1.6)	5.3 (1.7)	7.2 (1.9)	<i>F</i> = 75.97	<.001 ^{*,†,‡}	η_p^2 = .17
Boston Naming Test	8.5 (1.1)	4.9 (2.1)	8.0 (1.8)	<i>F</i> = 334.62	<.001 ^{*,†,‡,§}	η_p^2 = .47
Category Fluency	16.5 (4.6)	12.9 (4.0)	19.0 (5.2)	<i>F</i> = 117.85	<.001 ^{*,†,‡}	η_p^2 = .24
ADAS-Cog Immediate Word Recall (mean correct)	4.8 (1.0)	4.4 (1.2)	6.5 (0.8)	<i>F</i> = 260.98	<.001 ^{*,†,‡}	η_p^2 = .41
ADAS-Cog Delayed Word Recall (correct)	2.7 (1.5)	2.8 (2.0)	5.9 (1.5)	<i>F</i> = 274.13	<.001 ^{†,‡}	η_p^2 = .42
Arm of trial						
% in donepezil arm	34.0%	31.9%	32.3%	χ^2 = 1.81	.77	ϕ_c = .04
% in vitamin E arm	33.2%	35.6%	31.0%			
% in placebo arm	32.8%	32.5%	36.7%			

Abbreviations: aMCI, amnesic mild cognitive impairment; ADAS-Cog, Alzheimer's Disease Assessment Scale-Cognitive Subscale; ADL, activities of daily living; *APOE*, apolipoprotein E; AD, Alzheimer's disease; CDR, Clinical Dementia Rating; LM, Logical Memory; MMSE, Mini-Mental Status Examination. NOTE. Data are summarized as mean (standard deviation), unless otherwise indicated.

*Multi-domain aMCI significantly worse than (or fewer men than) single-Domain aMCI ($P < .01$).

[†]Single-domain aMCI significantly worse than (or older than) "false-positive" MCI ($P < .01$).

[‡]Multi-domain aMCI significantly worse than (or older than) "false-positive" MCI ($P < .01$).

[§]Single-domain aMCI significantly better than "false-positive" MCI ($P < .01$).

^{||}LM II Recall scores were available for a subset of the sample ($n = 617$).

E group (143/356; 40.2%) after 36 months, $\chi^2(1) = 5.32$, $P = .02$, $\phi_c = .10$. A Cox proportional hazards analysis showed that the donepezil group showed a significantly lower probability of progression from MCI to AD (hazard ratio = 0.71; 95% confidence interval = 0.52–0.98; $P = .04$; see Fig. 3).

Neuropsychological test data showed that the donepezil group had higher ADAS-Cog total scores relative to the placebo/vitamin E group at their 12-month ($P = .002$, $d = .34$) visit; no differences were seen at 24 or 36 months. On the word-list learning measure, the donepezil group recalled more words than the placebo/vitamin E group at the 12 ($P < .001$), 24 ($P = .02$), and 36-month ($P = .02$) visits (Table 3 and Fig. 2C). The donepezil group also showed better immediate recall on the paragraph-learning test relative to the placebo/vitamin E group at 12 ($P = .02$) and 24 ($P = .04$) months (Table 3 and Fig. 2D). There were no significant group by *APOE* ϵ 4 status interactions on ADAS-Cog total score or immediate memory performance. There were also no group differences on delayed recall of the words or the paragraph. Removal of the FP MCI group from the sample did not affect the results for any of the other neuropsychological measures included in the original study [3].

3.3. Comparison of aMCI-sd versus aMCI-md

The aMCI-md group was less educated and had a smaller percentage of men than the aMCI-sd group (Table 1). The groups did not differ in age. A chi-square analysis showed that the rate of progression to possible or probable AD after 36 months was significantly higher in the aMCI-md group (38.0%) than in the aMCI-sd group (35.3%), $\chi^2(1) = 77.00$, $P < .001$.

On the ADAS-Cog total score, the aMCI-md group performed worse than the aMCI-sd group at baseline ($P < .001$, $d = 0.37$), 12 ($P < .001$, $d = .35$), 24 ($P = .001$, $d = .35$), and 36 ($P < .001$, $d = .46$) months. The aMCI-md group performed worse than the aMCI-sd group on the word-list learning task at baseline ($P < .001$, $d = .33$), 12 ($P = .001$, $d = .32$), and 36 months ($P = .02$, $d = .27$), and worse on the paragraph-learning task at baseline ($P < .001$, $d = .31$), 12 ($P < .001$, $d = .47$), 24 ($P < .001$, $d = .39$), and 36 ($P < .001$, $d = .48$) months. The aMCI-md group also performed worse than the aMCI-sd group on delayed recall of the paragraph at 12 months ($P < .001$, $d = .39$) (but not other time points). The groups did not differ on delayed recall of words at any time point. There were no significant group

Table 2
Scores for the donepezil and vitamin E/placebo groups for the original MCI sample

	N (donepezil; vitamin E/placebo)	Donepezil: mean (SD)	Vitamin E/placebo: mean (SD)	F or χ^2	P value	Effect size
Demographics						
Age (years)	247; 509	73.2 (7.1)	73.0 (7.3)	$F = .10$.76	$d = .02$
Education (years)	247; 509	14.5 (2.9)	14.7 (3.2)	$F = .51$.47	$d = .06$
Gender (% male)	247; 509	56.3%	53.4%	$\chi^2 = .54$.46	$\phi_c = .03$
ADAS-Cog Immediate Word Recall (number of words correct; mean of three learning trials)						
Baseline	247; 509	5.2 (1.4)	5.1 (1.3)	$F = .45$.50	$d = .05$
12 months	181; 421	5.6 (1.4)	5.2 (1.6)	$F = 8.11$.005	$d = .26$
24 months	156; 347	5.5 (1.8)	5.2 (1.7)	$F = 3.28$.07	$d = .17$
36 months	139; 331	5.3 (1.9)	5.0 (1.9)	$F = 2.87$.09	$d = .17$
NYU Paragraph Immediate Recall (number of items correct)						
Baseline	247; 509	4.1 (2.3)	4.1 (2.3)	$F = .11$.75	$d = .03$
12 months	181; 420	4.0 (2.2)	3.6 (2.4)	$F = 2.46$.12	$d = .14$
24 months	157; 347	4.1 (2.1)	3.9 (2.3)	$F = .63$.43	$d = .08$
36 months	139; 330	3.7 (2.0)	3.7 (2.3)	$F < .01$.94	$d < .01$

Abbreviations: ADAS-Cog, Alzheimer's Disease Assessment Scale-Cognitive Subscale; NYU, New York University.
NOTE. Significant findings are in bold text.

by APOE $\epsilon 4$ status interaction on cognitive performance at any time point. There were also no significant group by treatment arm interactions on memory performance, indicating that donepezil was not more beneficial for one of the aMCI subgroups.

4. Discussion

Although diagnosed with MCI using conventional diagnostic criteria (as operationalized for multicenter studies), approximately one-third of participants in the ADCS donepezil trial in MCI [3] were identified as cognitively normal

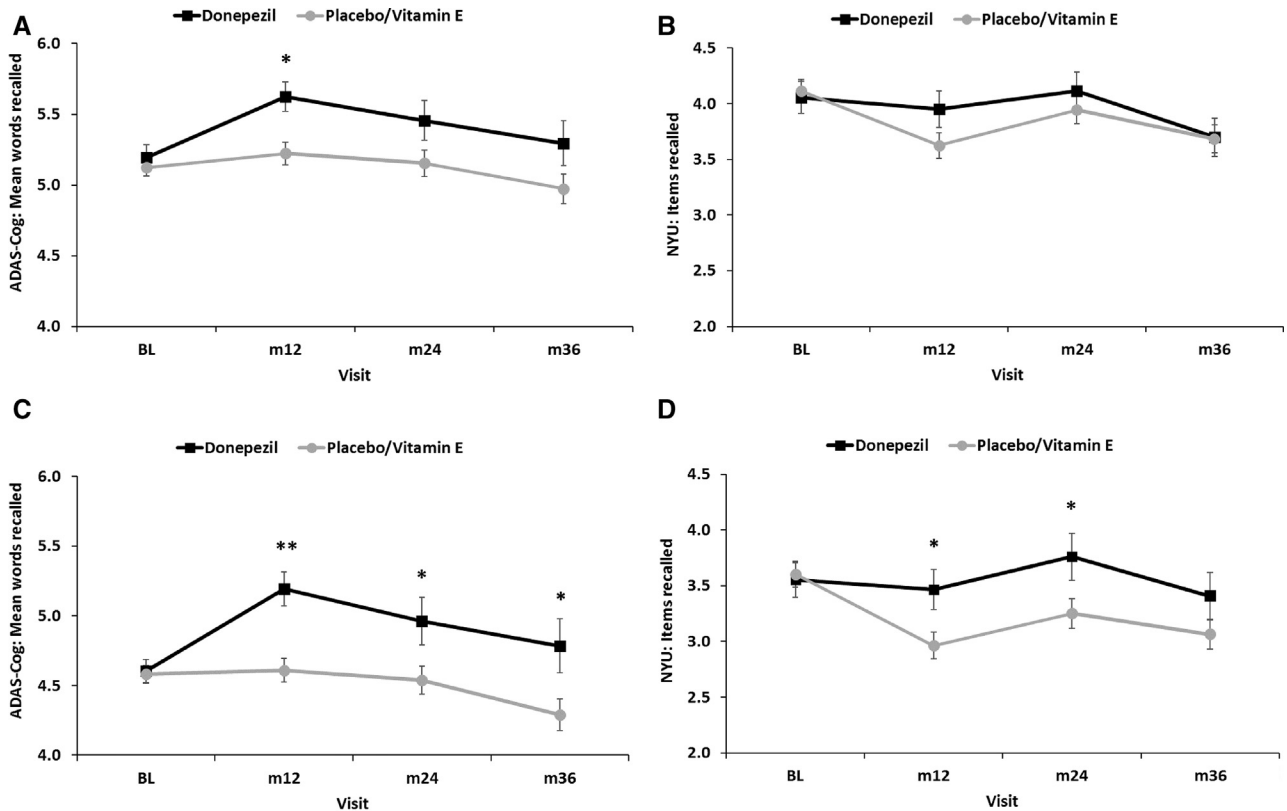


Fig. 2. Performance of the donepezil and placebo/vitamin E groups in the original MCI sample on the (A) ADAS-Cog Immediate Word Recall and the (B) NYU Paragraph Immediate Recall. Performance of the donepezil and placebo/vitamin E groups in the new MCI sample on the (C) ADAS-Cog Immediate Word Recall and the (D) NYU Paragraph Immediate Recall. * $P < .05$; ** $P < .001$. Abbreviations: ADAS-Cog, Alzheimer's Disease Assessment Scale-Cognitive Subscale; MCI, mild cognitive impairment; NYU, New York University.

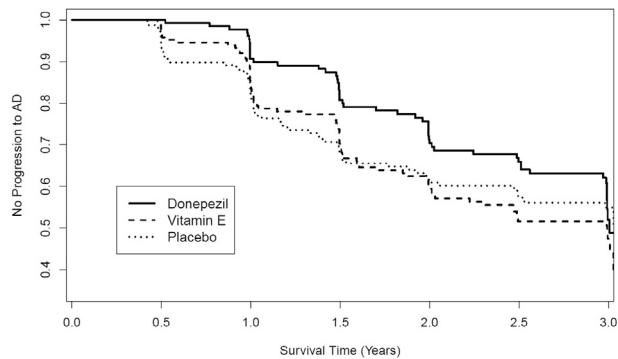


Fig. 3. Kaplan-Meier survival estimates showing the rate of progression from MCI to AD in the donepezil, vitamin E, and placebo arms in the *new MCI sample* that remained once the “false-positive” participants were removed. Abbreviations: AD, Alzheimer’s disease; MCI, mild cognitive impairment.

at baseline based on actuarial methods applied to multiple cognitive tests. This is consistent with the high rate of FP errors we have observed using these methods in a community-based sample [7] and in the ADNI data set [8]. It could be argued that these individuals simply represent “mild” or “early” forms of MCI rather than FP diagnostic errors. However, their scores on word-list learning and memory tasks were in the “low-average” range, and z-scores indexed to normative data were above the typical -1.5 SD cutoff to determine cognitive “impairment” [1,11] and even above a less stringent -1.0 SD cutoff [12,13]. Additionally, our prior analyses of ADNI data have shown that our actuarial “early MCI” and “late MCI” groups do not correspond with ADNI’s early versus late MCI designations [10]. This argues against the notion that our methods simply excluded the mildest MCI participants. Furthermore, if giving donepezil early in the course of AD maximizes the potential to observe treatment effects, removal of the mildest “true-positive” MCI cases would likely have diminished, rather than enhanced, a donepezil treatment effect.

Once FP MCI participants were removed, donepezil treatment was associated with a significantly lower rate of progression to AD dementia over 36 months (the pre-trial designated primary outcome) relative to placebo or treatment with vitamin E. Without FP MCI participants, donepezil treatment was also associated with significantly better ADAS-Cog total scores and better immediate memory performance relative to placebo or vitamin E throughout the 36-month trial. These beneficial effects of donepezil were only apparent in the first 12 months in the original analysis [3]. Our findings suggest that donepezil (and possibly other cholinesterase inhibitors) may be an effective treatment for MCI, but its benefits have been masked in clinical trials that included nonaffected, cognitively normal individuals in the MCI cohort. Better cognitive characterization of MCI samples before enrollment in clinical trials is likely to produce more promising results and will lead to more efficient trial designs with substantial cost savings.

Donepezil treatment had no apparent benefit relative to placebo or vitamin E on measures of delayed recall in either the original or new participant sample. This may be due to floor effects in the performance of those with MCI on delayed memory measures. In addition, it is possible that, as a cholinesterase inhibitor, donepezil exerts its effects primarily through improving attention/immediate memory [14]. Consistent with this possibility, a 24-week trial of donepezil in MCI showed no difference between donepezil and placebo groups in a primary analysis of a delayed verbal memory test, but there was a donepezil-related improvement in attention and psychomotor speed in a subsample of participants [15]. Other studies that examined donepezil in MCI and AD used global measures to evaluate cognition (e.g., ADAS-Cog total score, MMSE, Clinical Dementia Rating–Sum of Boxes) [16,17], making it difficult to determine the effect of donepezil on individual memory processes.

As originally conceived, MCI is a clinical diagnosis with no specific cutoffs or tests for determining impairment [1].

Table 3

Scores for the donepezil and vitamin E/placebo groups for the *new MCI sample with the “false-positive” participants removed*

	N (donepezil; vitamin E/placebo)	Donepezil: mean (SD)	Vitamin E/placebo: mean (SD)	<i>F</i> or χ^2	<i>P</i> value	Effect size
Demographics						
Age (years)	174; 356	73.6 (7.2)	73.7 (6.9)	<i>F</i> = .01	.94	<i>d</i> = .01
Education (years)	174; 356	14.2 (3.1)	14.2 (3.2)	<i>F</i> < .01	.98	<i>d</i> < .01
Gender (% male)	174; 356	56.3%	53.7%	χ^2 = .34	.56	ϕ_c = .03
ADAS-Cog Immediate Word Recall (number of words correct; mean of three learning trials)						
Baseline	174; 356	4.6 (1.1)	4.6 (1.1)	<i>F</i> = .03	.86	<i>d</i> = .02
12 months	124; 295	5.2 (1.3)	4.6 (1.4)	<i>F</i> = 14.79	<.001	<i>d</i> = .42
24 months	107; 239	5.0 (1.8)	4.5 (1.5)	<i>F</i> = 5.11	.02	<i>d</i> = .26
36 months	93; 221	4.8 (1.9)	4.3 (1.7)	<i>F</i> = 5.38	.02	<i>d</i> = .28
NYU Paragraph Immediate Recall (number of items correct)						
Baseline	174; 356	3.6 (2.0)	3.6 (2.1)	<i>F</i> = .07	.79	<i>d</i> = .03
12 months	124; 294	3.5 (2.0)	3.0 (2.1)	<i>F</i> = 5.26	.02	<i>d</i> = .25
24 months	108; 239	3.8 (2.2)	3.3 (2.1)	<i>F</i> = 4.36	.04	<i>d</i> = .24
36 months	93; 220	3.4 (2.0)	3.1 (1.9)	<i>F</i> = 2.01	.16	<i>d</i> = .17

Abbreviations: ADAS-Cog, Alzheimer’s Disease Assessment Scale-Cognitive Subscale; NYU, New York University.

NOTE. Significant findings are in bold text.

Clinical trials and research studies, however, require the use of cutoffs (e.g., -1.5 SD) to operationalize the MCI criteria and make them more applicable and standardized across multiple sites. Even then, a recent study showed that differences in the implementation of operationally defined conventional MCI criteria across the placebo arms of seven clinical trials led to significant variability in performance of MCI participants and in rates of progression [18].

Although the limitations we have observed in the diagnostic criteria may be specific to the context of a clinical trial or multi-site research study, there is growing evidence that comprehensive neuropsychological testing with actuarial decision-making can improve diagnostic accuracy for MCI [19–21]. We have proposed actuarial neuropsychological criteria for MCI that balance sensitivity to mild impairment (impairment defined as below -1 SD as opposed to -1.5 SD) with reliability (requires two impaired scores within a cognitive domain as opposed to a single impaired score) [12,20]. When compared with conventional diagnostic criteria for MCI (as operationalized for multicenter studies and clinical trials), this approach provides an MCI diagnosis that is more accurate and stable and more strongly associated with AD biomarkers and progression to AD [19,20]. Improved diagnostic accuracy could have a major impact on the design of future prospective studies of genetics, biomarkers, and treatment in MCI.

Several limitations of the present study should be noted. First, the neuropsychological diagnostic approach we employed was not flawless; 13 of the 226 participants identified as FP MCI progressed to possible or probable AD, indicating that they were indeed in a “mild” phase of MCI. However, this modest cost in sensitivity may be a small price to pay for the significant improvement in specificity that was gained (i.e., identifying 213 FP MCI diagnoses). Second, AD biomarkers were not collected in the original ADCS study, so we were unable to verify that the FP MCI participants we identified do not have significant underlying AD pathology. Furthermore, additional follow-up beyond the 36 months of the study was not obtained, so we cannot be certain that they did not progress to AD dementia at some point after the trial. Third, ADAS-Cog memory scores were used to classify the MCI groups (at baseline) and as an outcome measure (at subsequent visits). Although it would be ideal to have separate memory measures for these purposes, we were limited by the relatively brief neuropsychological test battery that was administered in the original trial. Although this raises some concern about circularity, it is important to note that the aMCI-sd, aMCI-md, and FP MCI groups did not differ on baseline ADAS-Cog memory scores and that a similar pattern of results was obtained with the NYU Paragraph Recall Test.

Overall our results suggest that donepezil has a beneficial effect on cognitive function and reduces the rate of progression to AD dementia in patients with MCI, but this effect is attenuated in clinical trials that include sizeable numbers of participants with a FP diagnosis of MCI. More emphasis on

comprehensive neuropsychological test data and actuarial methods when diagnosing MCI for clinical trials may increase trial efficiency and enhance the ability to discover significant drug effects.

Acknowledgments

Original data collection was funded by the Alzheimer's Disease Cooperative Study (NIA U01 AG10483), Pfizer, and Eisai. New data analysis was supported by National Institutes of Health grants R01 AG049810 (M.W.B.), K24 AG026431 (M.W.B.), P50 AG05131 (D.R.G., D.P.S., and S.D.E.), and the U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service (Career Development Award-2 1IK2 CX001415-01A1 to E.C.E.). D.R.G. serves as the editor for *Alzheimer's Research and Therapy* and as a paid consultant on Data Safety Monitoring Boards for Pfizer, Inc.; Elan, Inc.; and Balance Pharmaceuticals, Inc. D.P.S. serves as a consultant for Bristol-Myers Squibb. M.W.B. serves as a consultant for Novartis and receives royalties from Oxford University Press. The authors would like to thank Dr. Ron Petersen for reviewing a draft of the manuscript.

RESEARCH IN CONTEXT

1. Systematic review: The authors searched PubMed for studies related to diagnosis of mild cognitive impairment (MCI) for clinical trials. Results revealed that MCI is routinely diagnosed based on subjective complaints, an impaired score on a single objective memory test, clinical judgment of cognitive decline, and intact functional abilities. However, recent research suggests that this diagnostic approach may be overinclusive.
2. Interpretation: Our results support previous findings showing high rates of diagnostic errors based on conventional criteria, as approximately one-third of MCI participants in the Alzheimer's Disease Cooperative Study donepezil trial were identified as cognitively normal based on actuarial methods applied to multiple cognitive tests. Removal of “false-positives” unmasked beneficial effects of donepezil on cognition and rate of progression to Alzheimer's disease.
3. Future directions: Future research should emphasize comprehensive neuropsychological test data and actuarial methods when diagnosing MCI for clinical trials, as this may increase trial efficiency and enhance the ability to discover significant drug effects.

References

- [1] Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med* 2004;256:183–94.
- [2] Petersen RC, Morris JC. Mild cognitive impairment as a clinical entity and treatment target. *Arch Neurol* 2005;62:1160–3.
- [3] Petersen RC, Thomas RG, Grundman M, Bennett D, Doody R, Ferris S, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med* 2005;352:2379–88.
- [4] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 2010;74:201–9.
- [5] Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on mild cognitive impairment. *J Intern Med* 2004;256:240–6.
- [6] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:270–9.
- [7] Clark LR, Delano-Wood L, Libon DJ, McDonald CR, Nation DA, Bangen KJ, et al. Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? *J Int Neuropsychol Soc* 2013;19:635–45.
- [8] Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimers Dement* 2015;11:415–24.
- [9] Bangen KJ, Clark AL, Werhane M, Edmonds EC, Nation DA, Evangelista N, et al. Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE varepsilon4 genotype. *J Alzheimer Dis* 2016; 52:849–61.
- [10] Edmonds EC, Eppig J, Bondi MW, Leyden KM, Goodwin B, Delano-Wood L, et al. Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. *Neurology* 2016; 87:2108–16.
- [11] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56:303–8.
- [12] Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry* 2009;17:368–75.
- [13] Busse A, Hensel A, Guhne U, Angermeyer MC, Riedel-Heller SG. Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology* 2006;67:2176–85.
- [14] Klinkenberg I, Sambeth A, Blokland A. Acetylcholine and attention. *Behav Brain Res* 2011;221:430–42.
- [15] Salloway S, Ferris S, Kluger A, Goldman R, Griesing T, Kumar D, et al. Efficacy of donepezil in mild cognitive impairment: a randomized placebo-controlled trial. *Neurology* 2004;63:651–7.
- [16] Doody RS, Ferris SH, Salloway S, Sun Y, Goldman R, Watkins WE, et al. Donepezil treatment of patients with MCI: a 48-week randomized, placebo-controlled trial. *Neurology* 2009;72:1555–61.
- [17] Tan CC, Yu JT, Wang HF, Tan MS, Meng XF, Wang C, et al. Efficacy and safety of donepezil, galantamine, rivastigmine, and memantine for the treatment of Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimer Dis* 2014;41:615–31.
- [18] Petersen RC, Thomas RG, Aisen PS, Mohs RC, Carrillo MC, Albert MS. Randomized controlled trials in mild cognitive impairment: Sources of variability. *Neurology* 2017;88:1751–8.
- [19] Jak AJ, Urban S, McCauley A, Bangen KJ, Delano-Wood L, Corey-Bloom J, et al. Profile of hippocampal volumes and stroke risk varies by neuropsychological definition of mild cognitive impairment. *J Int Neuropsychol Soc* 2009;15:890–7.
- [20] Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimer Dis* 2014; 42:275–89.
- [21] Saxton J, Snitz BE, Lopez OL, Ives DG, Dunn LO, Fitzpatrick A, et al. Functional and cognitive criteria produce different rates of mild cognitive impairment and conversion to dementia. *J Neurol Neurosurg Psychiatry* 2009;80:737–43.