Efficacy of Cholinesterase Inhibitors in the Treatment of Neuropsychiatric Symptoms and Functional Impairment in Alzheimer Disease: A Meta-analysis

Nhi-Ha Trinh, MD, MPH
Jennifer Hoblyn, MB, MRCPsych, MPH
Subhanjoy Mohanty, PhD
Kristine Yaffe, MD

A LZHEIMER DISEASE (AD) IS THE most common form of dementia, accounting for more than 50% of all dementia cases, and the number of affected individuals will likely quadruple over the next several decades.1 Research in the treatment of AD has focused on reducing cognitive decline with cholinesterase inhibitors (ChIs), the primary class of medications currently available for this purpose. However, neuropsychiatric symptoms and functional deficits also contribute greatly to the disability associated with AD. As many as 80% of patients with AD will experience neuropsychiatric symptoms such as hallucinations, paranoia, agitation, and affective disturbances during the course of their illness.2,3 Functional impairment occurs in all patients with AD, first appearing as impairment in instrumental activities of daily living (IADL) tasks, including using the telephone and taking medications. As the disease progresses, alterations in basic activities of daily living (ADLs), such as feeding and dressing, become evident.4 Both neuropsychiatric symptoms and functional impairment contribute to an increase in caregiver burden, poor patient quality of life, and institutionalization.5,6 Cognitive deficits are linked to cholinergic dysfunction. More recently, Cholinesterase inhibitors are the primary treatment for the cognitive symptoms of Alzheimer disease (AD). Cholinergic dysfunction is also associated with neuropsychiatric and functional deficits, but results from randomized controlled trials of cholinesterase inhibitors are conflicting.

Context Cholinesterase inhibitors are the primary treatment for the cognitive symptoms of Alzheimer disease (AD). Cholinergic dysfunction is also associated with neuropsychiatric and functional deficits, but results from randomized controlled trials of cholinesterase inhibitors are conflicting.

Objective To conduct a systematic review and meta-analysis to quantify the efficacy of cholinesterase inhibitors for neuropsychiatric and functional outcomes in patients with mild to moderate AD.

Data Sources We performed a literature search of trials using MEDLINE (January 1966–December 2001), Dissertations Abstracts Online, PSYCHINFO, BIOSIS, PubMed, and the Cochrane Controlled Trials Register. We retrieved English- and non–English-language articles for review and collected references from bibliographies of reviews, original research articles, and other articles of interest. We searched for both published and unpublished trials, contacting researchers and pharmaceutical companies.

Study Selection We included 29 parallel-group or crossover randomized, double-blind, placebo-controlled trials of outpatients who were diagnosed as having mild to moderate probable AD and were treated for at least 1 month with a cholinesterase inhibitor. Sixteen trials included neuropsychiatric and 18 included functional measures.

Data Extraction Two investigators (N.H.T. and J.H.) independently extracted study methods, sources of bias, and outcomes. Neuropsychiatric outcomes were measured with the Neuropsychiatric Inventory (NPI, 0-120 points) and the Alzheimer Disease Assessment Scale, noncognitive (ADAS-noncog, 0-50 points) and were analyzed with the weighted mean difference method. Functional outcomes were measured with several activities of daily living (ADL) and instrumental activities of daily living (IADL) scales and analyzed with the standardized mean difference method.

Data Synthesis For neuropsychiatric outcomes, 10 trials included the ADAS-noncog and 6 included the NPI. Compared with placebo, patients randomized to cholinesterase inhibitors improved 1.72 points on the NPI (95% confidence interval [CI], 0.87-2.57 points), and 0.03 points on the ADAS-noncog (95% CI, 0.00-0.05 points). For functional outcomes, 14 trials used ADL and 13 trials used IADL scales. Compared with placebo, patients randomized to cholinesterase inhibitors improved 0.1 SDs on ADL scales (95% CI, 0.00-0.19 SDs), and 0.09 SDs on IADL scales (95% CI, 0.01 to 0.17 SDs). There was no difference in efficacy among various cholinesterase inhibitors.

Conclusions These results indicate that cholinesterase inhibitors have a modest beneficial impact on neuropsychiatric and functional outcomes for patients with AD. Future research should focus on how such improvements translate into long-term outcomes such as patient quality of life, institutionalization, and caregiver burden.

JAMA. 2003;289:210-216

©2003 American Medical Association. All rights reserved.
neuropsychiatric and functional deficits have also been associated with cholinergic dysfunction, with growing interest in ChIs for treatment of these symptoms in AD. Several clinical trials have included neuropsychiatric and functional measures, but have had conflicting results. In addition, these trials varied in numbers of subjects (ranging from 16 to 978) and used a variety of different measures of behavior and function. Thus, we conducted a systematic review and meta-analysis to quantify the efficacy of ChIs on neuropsychiatric and functional outcomes in patients with mild to moderate AD and living in the community.

METHODS

Identification of Trials
We searched MEDLINE (January 1966-December 2001), Dissertation Abstracts Online, PSYCHINFO, BIOSIS, PubMed, and the Cochrane Controlled Trials Register using the terms Alzheimer disease, dementia, and cholinesterase inhibitor. Based on the title of publication and abstract, we retrieved English-language and non-English-language articles for review, and also collected additional references from bibliographies of reviews, original research articles, and other articles of interest. We searched for both published and unpublished trials, contacting researchers and pharmaceutical companies.

Inclusion Criteria
Two investigators (N.H.T. and J.H.) independently reviewed all pertinent articles using predetermined inclusion criteria. A trial was included if (1) it was randomized, double-blinded, and placebo-controlled; (2) the design of the trial was either parallel or crossover; (3) if a crossover trial, it had a washout period greater than 1 week; (4) patients enrolled were outpatients diagnosed as having mild to moderate, “probable” AD according to National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria, with a baseline Mini-Mental State Examination score of 10 to 26; (5) the trial involved more than 1 month of treatment with 1 of the following ChIs: donepezil, rivastigmine, galantamine, metrifonate, physostigmine, rivastigmine, tacrine, tacrine with lecithin, and velnacrine; (6) neuropsychiatric outcomes were measured with the most common neuropsychiatric scales used, the non-cognitive, and placebo groups. For treatment and placebo groups, we calculated effect sizes for each outcome measure to compare treatment and placebo groups. For treatment and placebo groups in each trial, we collected the mean and variance for the change in scores from baseline, using the ITT analysis. If data from the ITT analysis were not available, we used the completed subjects analysis. For trials not providing the above data directly, we calculated the data from given information, using widely used statistical techniques. In dose-ranging trials with multiple treatment groups, we combined treatment groups by weighting the effects for each subgroup by its sample size.

For the NPI and the ADAS-noncog meta-analyses, we used the weighted mean difference method, in which effect size is the difference in the change of score from baseline between the treatment and placebo groups; we present results as points of the scale. If trials reported both scales, we used the NPI

<p>| Table 1. Trials of Cholinesterase Inhibitors Using the NPI and ADAS-noncog Scales for Neuropsychiatric Outcomes in Patients With Alzheimer Disease |
|------------------|------------------|------------------|------------------|------------------|</p>
<table>
<thead>
<tr>
<th>Source</th>
<th>Subjects, No.</th>
<th>Medication, mg/kg per d</th>
<th>Trial Length, d</th>
<th>Intention-to-Treat Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPI*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morris et al, 1998</td>
<td>408</td>
<td>Metrifonate, 0.65</td>
<td>168</td>
<td>Yes</td>
</tr>
<tr>
<td>Dubois et al, 1999</td>
<td>605</td>
<td>Donepezil, 10 mg/d</td>
<td>168</td>
<td>Yes</td>
</tr>
<tr>
<td>Raskind et al, 2000</td>
<td>264</td>
<td>Metrifonate, 50 mg/d</td>
<td>182</td>
<td>Yes</td>
</tr>
<tr>
<td>Tariot et al, 2000</td>
<td>978</td>
<td>Galanthamine, 8, 16, or 24 mg/d</td>
<td>150</td>
<td>Yes</td>
</tr>
<tr>
<td>Rockwood et al, 2001</td>
<td>386</td>
<td>Galanthamine, 24 or 32 mg/d</td>
<td>90</td>
<td>Yes</td>
</tr>
<tr>
<td>Winblad et al, 2001</td>
<td>286</td>
<td>Donepezil, 10 mg/d</td>
<td>365</td>
<td>Yes</td>
</tr>
<tr>
<td>ADAS-noncog</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis et al, 1992</td>
<td>215</td>
<td>Tacrine, 40 or 80</td>
<td>42</td>
<td>Yes</td>
</tr>
<tr>
<td>Farlow et al, 1992</td>
<td>468</td>
<td>Tacrine, 20, 40, or 80</td>
<td>84</td>
<td>Yes</td>
</tr>
<tr>
<td>Knapp et al, 1994</td>
<td>653</td>
<td>Tacrine, 80, 120, or 160</td>
<td>210</td>
<td>Yes</td>
</tr>
<tr>
<td>Wood, 1994</td>
<td>154</td>
<td>Tacrine, 80 mg/d (mean dosage)</td>
<td>84</td>
<td>No</td>
</tr>
<tr>
<td>Forette et al, 1995</td>
<td>130</td>
<td>Tacrine, 40 or 80</td>
<td>42</td>
<td>Yes</td>
</tr>
<tr>
<td>Becker et al, 1996</td>
<td>50</td>
<td>Metrifonate, 2.1 mg/kg per wk</td>
<td>84</td>
<td>No</td>
</tr>
<tr>
<td>Zemlan et al, 1996</td>
<td>309</td>
<td>Velnacrine, 30, 75, 150, or 225</td>
<td>42</td>
<td>Yes</td>
</tr>
<tr>
<td>Becker et al, 1998</td>
<td>47</td>
<td>Metrifonate, 2.9 mg/kg per wk</td>
<td>180</td>
<td>Yes</td>
</tr>
<tr>
<td>Jann et al, 1999</td>
<td>395</td>
<td>Metrifonate, 50 mg/d</td>
<td>42</td>
<td>No</td>
</tr>
<tr>
<td>Moller et al, 1999</td>
<td>181</td>
<td>Physostigmine patch, 30 or 60 mg (5.7 mg/d)</td>
<td>168</td>
<td>No</td>
</tr>
</tbody>
</table>

Abbreviations: ADAS-noncog, Alzheimer Disease Assessment Scale, noncognitive; NPI, Neuropsychiatric Inventory.

Three trials also reported the ADAS-noncog.

Statistical Analysis
Using a random-effects model and widely used meta-analytic methods, we calculated effect sizes for each outcome measure to compare treatment and placebo groups. For treatment and placebo groups in each trial, we collected the mean and variance for the change in scores from baseline, using the ITT analysis. If data from the ITT analysis were not available, we used the completed subjects analysis. For trials not providing the above data directly, we calculated the data from given information, using widely used statistical techniques. In dose-ranging trials with multiple treatment groups, we combined treatment groups by weighting the effects for each subgroup by its sample size.

For the NPI and the ADAS-noncog meta-analyses, we used the weighted mean difference method, in which effect size is the difference in the change of score from baseline between the treatment and placebo groups; we present results as points of the scale. If trials reported both scales, we used the NPI
results because the NPI scale includes more categories of neuropsychiatric problems, such as anxiety, apathy, and aggression. Because of diversity of scales reported for functional measures (4 scales in ADL and 2 scales in IADL domains), we used the standardized mean difference method to calculate effect size, defined as the difference in change of score from baseline between the treatment and placebo groups divided by the pooled SD of the trial; we present results in SDs.17

We present results as overall summary estimates combining all ChIs, together with 95% confidence intervals (CIs). For each outcome, we also conducted separate meta-analyses for each type of ChI if 3 or more trials existed for that agent. We tested for statistical homogeneity with the Mantel-Haenszel test using a small but statistically significant benefit from the use of ChIs

**RESULTS**

We examined a total of 3000 titles or abstracts in the literature search to identify published trials for possible inclusion. We identified 8 unpublished trials; none met our predetermined inclusion criteria. We reviewed 152 published trials in detail; of the 36 trials meeting inclusion criteria, 7 of these trials included overlapping data, leaving 29 trials. Of these, 27 were parallel-group trials, while 2 were crossover trials. Twenty-three of the eligible trials required that we calculate additional data from provided information: in 2 trials, we calculated the SD for the change in score from initial and final score SDs; in 3 trials, we calculated SDs from SEs; and in another 3 trials, we calculated SDs from CIs. For 15 trials, we calculated pooled SDs from P values; 10 trials gave exact P values, while 5 trials required us to estimate P values. For 14 trials that used more than 1 dose compared with a single placebo group, we combined the treatment groups by weighting the effects for each subgroup by its sample size.

**Neuropsychiatric Outcomes**

Sixteen of the eligible trials included neuropsychiatric outcomes. Six trials reported the NPI, and 10 the ADAS-noncog (Table 1). Three trials reported data from both scales; for these, we used the NPI in our analysis: 1 trial reported inadequate data for the ADAS-noncog, and 2 trials reported both the NPI and ADAS-noncog. Most (12 of 16) trials reported data from the ITT analysis, with 4 trials reporting only the completed subject analysis. Initially, our goal was to combine the NPI and ADAS-noncog scales in an overall summary estimate. Because of evidence of heterogeneity when we combined the scales (P<.01), we calculated separate meta-analyses for the NPI and ADAS-noncog outcomes.

For the 6 trials using the NPI, summary meta-analysis indicated that patients randomized to ChIs improved 1.72 points compared with placebo (95% CI, 0.87-2.57 points), indicating a small but statistically significant benefit from the use of ChIs (Figure 1A). For the 10 trials using the ADAS-noncog, summary meta-analysis indicated that patients randomized to ChIs improved 0.03 points compared with placebo (95% CI, 0.00-0.05 points), indicating a trend toward benefit in neuropsychiatric dysfunction from the use of ChIs (Figure 1B). For both the ADAS-noncog and the NPI meta-analyses, the tests of heterogeneity were not statistically significant (P=.99 for both).

**Functional Outcomes**

Fourteen eligible trials included ADL measures (Table 2). Eight trials re-
ported the ITT analyses and 6 the completed subjects analyses. Two of the included trials were crossover trials. For the fourteen trials using ADLs, summary meta-analysis indicated that patients randomized to ChIs improved by 0.10 SDs compared with placebo (95% CI, 0.00-0.19), indicating a trend in benefit in ADLs from the use of ChIs (FIGURE 2A). The test of heterogeneity was nonsignificant (P = .36).

Thirteen eligible trials included IADL outcomes. Ten trials reported data from the ITT analysis and 3 from the completed subjects analysis (Table 2). For the thirteen trials using IADLs, summary meta-analysis indicated that patients randomized to ChIs improved by 0.09 SDs compared with placebo (95% CI, 0.01-0.17), indicating a small but statistically significant benefit in IADLs from the use of ChIs (Figure 2B). A test of homogeneity was not significant (P = .47).

### Additional Analyses

To determine if our meta-analyses were sensitive to choices in methodology, we conducted a series of sensitivity analyses. First, for the 2 trials in which both the ADAS-noncog and NPI were used, we recalculated both meta-analyses by reassigning the 2 trials to the ADAS-noncog meta-analysis while omitting them from the NPI meta-analysis. There was no change in the magnitude or statistical significance of the results of either meta-analysis. Next, we determined whether the type of ChI had an effect on the results using subgroup analyses. For each outcome, if there were 3 or more trials using the same ChI, we calculated separate meta-analyses for each drug. We calculated a metrifonate subanalysis for the NPI scale, a tacrine subanalysis for the ADAS-noncog scale, a tacrine subanalysis for the ADL scale, and a tacrine and physostigmine subanalysis for the IADL scale. We found similar effect sizes but wider CIs for the subanalyses, with no clear difference with any one ChI. Finally, when we calculated separate subanalyses comparing the results of the ITT analyses with the completed subjects analyses for each outcome, we found similar results. The effect size and statistical significance were slightly higher for the completed subject NPI analyses (summary estimate, 1.96 points improvement; 95% CI, 1.07-2.84) and for the ADAS-noncog analyses (summary estimate, 0.39 points improvement; 95% CI, 0.30-0.47) but were not different for the functional outcomes.

For all 4 meta-analyses, funnel plots revealed no clear asymmetry to suggest the potential for publication bias against trials whose results showed no benefit of treatment over placebo. Although funnel plots themselves have low statistical power, additional formal testing using the Kendall tau showed no statistical evidence for publication bias in the meta-analyses of ADAS-noncog (P = .85), NPI (P = .84), ADL (P = .70), or IADL (P = .47).

### COMMENT

Based on meta-analyses of 29 trials (16 with neuropsychiatric outcomes and 18 with functional outcomes), our results support the hypothesis that ChIs...
have a modest beneficial role in treating neuropsychiatric symptoms and reducing functional impairment in patients with mild to moderate AD and living in the community. For neuropsychiatric outcomes, our analysis showed a small but statistically significant benefit of ChIs on the NPI scale and a trend in benefit on the ADAS-noncog scale. For functional outcomes, treatment with ChIs resulted in a small but statistically significant benefit of ChIs for IADLs and a trend in benefit for ADLs. Each of the ChIs had similar beneficial effects.

These results are consistent with a prior meta-analysis of 4 trials of tacrine in which there was a small but statistically significant benefit on the ADAS-noncog.

Our findings of a modest improvement in neuropsychiatric scales in patients treated with ChIs are important because neuropsychiatric problems are common in AD and are major contributors to loss of autonomy, morbidity, and need for nursing home placement. Although neuropsychological data are sparse, plausible biological mechanisms support the role of acetylcholine in modulating behavior in AD, and there has been increasing interest in ChIs for treatment of neuropsychiatric symptoms. Since the baseline ADAS-noncog scores of these trials ranged from 3.8 to 7.7, the improvement of less than 1 point on the ADAS-noncog scale is very small. However, the ADAS-noncog has been criticized for its lack of sensitivity, as it does not include some behaviors common in AD such as aggressiveness and anxiety. The NPI, which does include these neuropsychiatric domains, has been thought to be more sensitive to change.

An almost 2-point improvement on the NPI is equivalent to a decrease in frequency or severity of a particular neuropsychiatric symptom, and is notable since the baseline NPI scores of these trials ranged from 9.2 to 13.9.

Given the probable benefit ChIs may have on neuropsychiatric scales, our data suggest that for patients with mild to moderate AD who have neuropsychiatric disturbances, ChIs should be considered a therapeutic option. If neuropsychiatric symptoms are not improved after a trial of ChIs, then other agents such as antipsychotics or mood stabilizers should be considered. It is important to note that our observed effect size for neuropsychiatric outcomes measured with the NPI is similar to that reported in most randomized trials using other medications for neuropsychiatric symptoms in dementia, such as valproic acid and risperidol. As with all medication, however, the decision to use ChIs needs to be considered in light of adverse effects, cost, and feasibility.

As well as with neuropsychiatric symptoms, the presence of functional impairment can lead to poor health-related quality of life in patients with AD. The finding of a small improvement in ADLs and IADLs in patients treated with ChIs is of clinical importance since all patients with AD by definition have functional impairment, and functional decline increases with dis...
ease progression. To interpret the results of functional outcomes in SD units, some authors have argued that effect sizes of 0.20 in magnitude are clinically detectable\(^4\), thus, our results in the magnitude of 0.10 SD may be at the lower limit of clinical detectability. Based on prior longitudinal data of functional decline in AD, effect sizes in the magnitude of 0.10 SD would be similar to preventing a 2-months per year decline in a typical patient with AD.\(^6\) A recent study suggests that donepezil treatment for a year may have a beneficial effect on reducing functional impairment in patients with mild to moderate AD.\(^6\) However, because functional impairment is closely linked to cognitive impairment, it is unknown whether the benefit to functional impairment from ChIs is due to reducing cognitive impairment or an independent mechanism.

Some limitations to our meta-analyses restrict the interpretation of our results. First, not all subjects enrolled in the trials had neuropsychiatric problems, and the magnitude of the effect of ChIs for neuropsychiatric problems might be different for patients who all have difficult behaviors. Thus, more randomized clinical trials of ChIs are needed, specifically enrolling patients with AD and neuropsychiatric problems and focusing on sensitive neuropsychiatric measures. Second, although most of our included trials excluded patients receiving psychotropic medications, some of the trials allowed the use of medications such as short-acting benzodiazepines for sedation; the use of such medications may also have altered our results. In addition, we had limited power to perform separate meta-analyses by each ChI, and the trials included also varied in ChI treatment regimens and in duration of treatment. Despite these differences, subanalyses and tests of homogeneity indicated that the trial results were combinable. We also were unable to evaluate individual behaviors, such as aggression or paranoia, or functional abilities, such as the ability to cook, bathe, or dress, as most trials did not provide subscores. Finally, because neuropsychiatric and functional outcomes were secondary outcomes, data were often not fully reported; however, we used widely used meta-analytic techniques to calculate the non-reported data.

Despite these limitations, we believe our results are the first attempt to quantitatively synthesize the efficacy of a variety of ChIs for neuropsychiatric symptoms and functional impairment, and to suggest a modest benefit in neuropsychiatric outcomes and functional outcomes from ChIs. The results of our meta-analyses indicate the need for additional focused research on neuropsychiatric and functional outcomes in patients with AD. Exploring these issues will possibly broaden treatment for AD, an illness with growing importance to public health.

Author Affiliations: Department of Psychiatry, Massachusetts General Hospital, Boston (Dr Trinh); Veterans Affairs Palo Alto Health Care System and Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, Calif (Dr Hoblyn); Center for Astrophysics, Harvard University, Boston (Dr Mohanty); and Departments of Psychiatry, Neurology, and Epidemiology, University of California, San Francisco and the San Francisco Veterans Affairs Medical Center, San Francisco, Calif (Dr Yaffe).

Author Contributions: Study concept and design: Trinh, Hoblyn, Yaffe. Acquisition of data: Trinh, Hoblyn, Yaffe. Analysis and interpretation of data: Trinh, Hoblyn, Mohanty, Yaffe. Drafting of the manuscript: Trinh, Hoblyn, Yaffe. Critical revision of the manuscript for important intellectual content: Trinh, Hoblyn, Mohanty, Yaffe. Statistical expertise: Trinh, Hoblyn, Mohanty, Yaffe. Administrative, technical, or material support: Hoblyn. Study supervision: Yaffe.

Funding/Support: Dr Yaffe was supported by Public Health Service grant K23 AG00888 and is an American Federation for Aging Research Paul Beeson Faculty Scholar in Aging Research.

REFERENCES

5. Coen RF, Swanwick GRJ, O’Boyle CA, Coakley D. Behaviour disturbance and other predictors of carer burden in Alzheimer’s disease. Int J Geriatr Psychia-

©2003 American Medical Association. All rights reserved.
CHOLINESTERASE INHIBITORS IN ALZHEIMER DISEASE


