

# Efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer's disease: A systematic review and meta-analysis

Richard A Hansen<sup>1</sup>  
 Gerald Gartlehner<sup>2</sup>  
 Aaron P Webb<sup>1</sup>  
 Laura C Morgan<sup>3</sup>  
 Charity G Moore<sup>4</sup>  
 Daniel E Jonas<sup>3</sup>

<sup>1</sup>School of Pharmacy, <sup>3</sup>Cecil G Sheps Center for Health Services Research, University of North Carolina, Chapel Hill, NC, USA; <sup>2</sup>Department for Evidence-based Medicine and Clinical Epidemiology, Danube University, Krems, Austria; <sup>4</sup>Center for Research on Health Care Data Center; University of Pittsburgh, Pittsburgh, PA, USA

**Abstract:** Pharmacologic treatments for Alzheimer's disease include the cholinesterase inhibitors donepezil, galantamine, and rivastigmine. We reviewed their evidence by searching MEDLINE®, Embase, The Cochrane Library, and the International Pharmaceutical Abstracts from 1980 through 2007 (July) for placebo-controlled and comparative trials assessing cognition, function, behavior, global change, and safety. Thirty-three articles on 26 studies were included in the review. Meta-analyses of placebo-controlled data support the drugs' modest overall benefits for stabilizing or slowing decline in cognition, function, behavior, and clinical global change. Three open-label trials and one double-blind randomized trial directly compared donepezil with galantamine and rivastigmine. Results are conflicting; two studies suggest no differences in efficacy between compared drugs, while one study found donepezil to be more efficacious than galantamine, and one study found rivastigmine to be more efficacious than donepezil. Adjusted indirect comparison of placebo-controlled data did not find statistically significant differences among drugs with regard to cognition, but found the relative risk of global response to be better with donepezil and rivastigmine compared with galantamine (relative risk = 1.63 and 1.42, respectively). Indirect comparisons also favored donepezil over galantamine with regard to behavior. Across trials, the incidence of adverse events was generally lowest for donepezil and highest for rivastigmine.

**Keywords:** Alzheimer's disease, cholinesterase inhibitor, donepezil, galantamine, rivastigmine, systematic review, meta-analysis, indirect comparison

## Introduction

Alzheimer's disease is an age-associated neurodegenerative disorder, affecting approximately 24 million individuals worldwide (Hebert et al 2003). Primary manifestations of Alzheimer's disease include cognitive impairment, alterations in behavior, and reduced ability to perform activities of daily living. Nonpharmacologic and pharmacologic interventions are available, although none prevents or cures the disease. Non-pharmacologic interventions primarily address behavioral disturbances (eg, task simplification, environmental modification, minimal excess stimulation, etc) and other sources of cognitive impairment (eg, treating comorbid medical conditions, minimizing or eliminating drugs with deleterious cognitive side effects) (Cummings et al 2002). Pharmacologic therapies are intended to slow the progression of disease and improve symptoms. Drugs currently approved for Alzheimer's include cholinesterase inhibitors (donepezil hydrochloride [donepezil], galantamine hydrochloride [galantamine], rivastigmine tartrate [rivastigmine], and tacrine hydrochloride [tacrine]) and memantine, an N-methyl-D-aspartate (NMDA) receptor antagonist.

Currently available drugs have demonstrated modest benefits, although their place in the treatment of Alzheimer's disease has been heavily debated. For example,

Correspondence: Richard A Hansen  
 School of Pharmacy, Division  
 of Pharmaceutical Outcomes and Policy,  
 University of North Carolina at Chapel  
 Hill, 2205 Kerr Hall; Campus Box 7360,  
 Chapel Hill, NC 27599, USA  
 Tel +1 919 966 7517  
 Fax +1 919 966 8486  
 Email rahansen@unc.edu

the American Psychiatric Association (2007) recommends the cholinesterase inhibitors donepezil, galantamine, and rivastigmine for mild to moderate Alzheimer's disease, and suggests that they may be helpful for patients with severe disease. Memantine—a drug labeled for moderate to severe disease—is recommended for moderate to severe disease. In contrast, the National Institute for Clinical Excellence (NICE 2007), an organization responsible for providing guidance to the UK's National Health Service, only recommends donepezil, galantamine, and rivastigmine as options for the treatment of moderate Alzheimer's disease. Memantine is not recommended unless it is being used as part of a clinical trial. Although NICE takes a relatively aggressive stance in comparison to other organizations influencing payment policy, the high cost and modest benefits of these drugs continue to raise concerns.

To date, numerous review articles have been published that summarize the clinical efficacy and safety of drugs for the treatment of Alzheimer's disease (Geldmacher 2003, 2007; Lanctot et al 2003a, 2003b; Trinh et al 2003; Masterman 2004; Ritchie et al 2004; Forchetti 2005; Harry and Zakzanis 2005; Kadoszkiewicz et al 2005; Birks 2006; Loveman et al 2006; Loy and Schneider 2006; Schmitt et al 2006; Takeda et al 2006; Beier 2007; Hansen et al 2007). Most reviews have focused on the second-generation cholinesterase inhibitors (ie, donepezil, galantamine, and rivastigmine), since they have largely supplanted the first approved drug in this class (ie, tacrine) and are pharmacologically unique from memantine—a drug that targets glutamate rather than acetylcholine and has been studied primarily in more severe disease. These review articles can be differentiated by which specific drugs were included, the types of outcomes that were assessed, and by whether the focus was on overall efficacy (eg, placebo-controlled trials) or on comparative efficacy (eg, head-to-head trials or indirect comparison using placebo-controlled trials). A number of good-quality reviews have synthesized evidence regarding the overall efficacy of donepezil, galantamine, and rivastigmine, although most focus exclusively on specific outcome domains rather than a broad spectrum of outcome measures. Additionally, reviews synthesizing comparative evidence are sparse, in large part because of the quality and quantity of head-to-head trials (Wilkinson et al 2002; Wilcock et al 2003; Jones et al 2004; Bullock et al 2005). Existing head-to-head evidence cannot be pooled because of significant differences in trial populations and design, and only one review has attempted to make indirect comparisons using placebo-controlled data (Harry and Zakzanis 2005). Several other systematic reviews and

meta-analyses have narratively compared effect sizes across drugs, acknowledging the potential limitations of making unadjusted indirect comparisons.

We conducted a systematic review and meta-analysis of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer's disease. We attempted to elaborate on previous review articles by including a broad spectrum of outcome measures (ie, cognition, function, behavior, and global assessment), and emphasizing comparative evidence. We made adjusted indirect comparisons using placebo-controlled data for outcome measures with sufficient data.

## Methods

### Literature search

We searched MEDLINE®, Embase, The Cochrane Library, and the International Pharmaceutical Abstracts for studies addressing the general or comparative effectiveness of donepezil, galantamine, or rivastigmine for Alzheimer's disease. Sources were searched from 1980 to 2007 (July) to identify literature relevant to the scope of our topic. We manually searched reference lists of relevant review articles and letters to the editor. Additionally, we hand-searched the US Center for Drug Evaluation and Research database and the National Institutes of Health clinical trials registry ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)) to identify unpublished research.

### Inclusion criteria

Results from randomized, controlled trials comparing one cholinesterase inhibitor to another or to placebo were included. Community dwelling and nursing home populations were eligible. Trials had to last at least 12 weeks and include at least one measure reflecting the following: cognition, function, behavior, or clinical global assessment of change. Studies with statistically significant baseline differences between treatment groups that were deemed to affect outcomes were excluded, as were studies with other fatal flaws in study design or data analysis that contributed to a "poor" quality rating for internal validity. Comparative trials were not required to be double-blinded because a priori we knew that the majority of evidence comes from open-label trials. Placebo-controlled trials were required to be double-blinded.

### Quality assessment

We assessed the internal validity (quality) of trials based on predefined criteria developed by the US Preventive Services Task Force (ratings: good, fair, or poor) (Harris et al 2001) and the National Health Service Centre for Reviews

and Dissemination (2001). Elements of internal validity assessment included, among others, randomization and allocation concealment, similarity of compared groups at baseline, use of intention-to-treat (ITT) analysis, and overall and differential loss to follow-up. Two independent reviewers assigned quality ratings; they resolved any disagreements by discussion and consensus or by consulting a third independent party. Trials that had a fatal flaw in one or more categories were rated as poor quality and excluded from this analysis.

## Data abstraction

Trained reviewers abstracted data from each study and assigned an initial quality rating; a senior reviewer read each abstracted article, evaluated the completeness of the data abstraction, and confirmed the quality rating. We abstracted the following data from included trials: study design, eligibility criteria, intervention (drugs, dose, duration), additional medications allowed, methods of outcome assessment, population characteristics, sample size, loss to follow-up, withdrawals attributed to adverse events, results, and adverse events reported. We recorded ITT results if available.

## Outcome measures

Measurement scales varied across studies. We grouped measurement scales according to the general domain being assessed: cognition, function, behavior, and global assessment of change. We tried to limit outcome measures to a single measurement scale within each domain, although for some domains (eg, function), no single scale was used in the majority of trials so we abstracted data from most commonly used scales.

### Cognition

We focused on the Alzheimer's Disease Assessment Scale-Cognitive section (ADAS-cog) as the primary measure of cognition (Rosen et al 1984). Higher scores on this 11-question, 70-point scale reflect more severe cognitive deficits. Data were coded as the mean and standard deviation of the change from baseline to endpoint.

### Function

Because measures of function are particularly variable among clinical trials, we included all of the following: the Alzheimer's Disease Cooperative Studies Activities of Daily Living Inventory (ADCS/ADL); the Alzheimer's Disease Functional Assessment and Change Scale (ADFACS); the Bristol Activities of Daily Living Scale (BADLS); the Caregiver-rated Modified Crichton Scale (CMCS); the

Disability Assessment for Dementia (DAD); the Interview for Deterioration in Daily living activities in Dementia (IDDD); the Nurses Observation Scale for Geriatric Patients Activities of Daily Living subscale (NOSGER-IADL); and the Progressive Deterioration Scale (PDS). Functional outcome measures were initially coded as the mean and standard deviation of the mean change from baseline to endpoint for each measure, and later converted to a standardized effect size (Hansen et al 2007).

### Behavior

Behavioral outcomes were limited to the Neuropsychiatric Inventory (NPI) (Cummings et al 1994), a 144-point scale, with higher scores reflecting greater severity. Data were coded as the mean and standard deviation of the change from baseline to endpoint.

### Global assessment of change

The Clinician Interview-Based Impression of Change Incorporating Caregiver Information (CIBIC+) scale was recorded as the primary global assessment of change (Knopman et al 1994). The CIBIC+ includes a 7-point Likert scale to code the overall impression of change ("7" marked worsening; "4" no change; "1" marked improvement). This scale was coded as a binary outcome to classify responders (<4) and nonresponders ( $\geq 4$ ). The Clinical Global Impression of Change (CGI-C) was included as a secondary measure of global assessment of change (Schneider et al 1997). The CGI-C reflects the same 7-point Likert scale as the CIBIC+, although it does not follow a semi-structured format with caregiver input. Both scales were coded as the number of responders and non-responders among participants randomized to each treatment.

## Data synthesis

Head-to-head studies were described, but not quantitatively combined because there were too few studies and the majority were open-label rather than double-blinded. Placebo-controlled data were combined in meta-analysis for each outcome measure. For continuous data collected using the same measurement scale (eg, cognition and behavior), we conducted an analysis of the weighted mean difference. The weighted mean difference reflects the difference in change from baseline to endpoint for active treatment compared with placebo, weighted by the inverse variance (ie, studies with smaller variance, and likely larger sample size, given more weight). For functional outcomes, which were assessed on a number of different measurement scales, we calculated

a standardized mean difference (ie, standardized effect size). The standardized mean difference, sometimes referred to as *d* (Cohen 1988), is a scale-free measure of the separation between two group means. A standardized effect size of “0” is comparable with no difference between active treatment and placebo. Global assessment of change was analyzed as the relative risk of being classified as a responder for treatment compared with placebo. Our primary analysis was limited to trials reporting the CIBIC+, although sensitivity analyses pooled data for the CIBIC+ and CGI-C.

All meta-analyses specified a random effects model, which assumes that variability in effect sizes is due to sampling error plus unique differences in the set of true population effect sizes. We tested for heterogeneity of treatment effects using the  $I^2$  statistic (Higgins et al 2003). To estimate possible publication bias caused by the tendency of published studies to be positive, we used funnel plots (Egger et al 1997).

Because no head-to-head evidence was available for the majority of drug comparisons, we conducted adjusted indirect comparisons of placebo-controlled trials employing the method proposed by Bucher and colleagues (1997). Adjusted indirect comparisons assess the relative benefits of two treatments when they have not been compared directly with each other, but have each been evaluated relative to a common comparator (Glenny et al 2005). Evidence suggests that indirect comparisons agree with head-to-head trials if component studies are similar and treatment effects are expected to be consistent in patients included in different trials. For indirect comparisons of outcomes reflecting continuous data (eg, weighted mean difference), our reported values can be interpreted as the pooled weighted mean difference for *Drug A* minus the pooled weighted mean difference for *Drug B*. Values close to zero reflect no differences between compared drugs. For binary data (eg, relative risk of global response), our reported values can be interpreted as the relative risk of responding with *Drug B* compared with placebo over the relative risk of responding with *Drug A* compared with placebo. Thus, overall relative risk values less than 1.0 favor *Drug A*, while relative risk values greater than 1.0 favor *Drug B*.

For completeness in assessing the benefits and risks of these drugs, we reviewed adverse events. Data from included trials were abstracted, and the mean incidence and 95% confidence intervals (CI) for specific adverse events were calculated. The number of withdrawals, and the number of withdrawals due to adverse events, were recorded and summarized by drug. Meta-analysis was used to quantify

the relative risk of withdrawing for each drug compared with placebo.

## Results

We found 1,476 unduplicated citations (Appendix 1). Of these, 1,112 citations were excluded after reviewing the abstract and 321 full-text articles were retrieved. After full-text review, 166 citations were excluded for failure to meet eligibility criteria, and 2 for poor methodological quality; 120 citations were relevant for background information, and 33 articles on 26 studies were included in the review. A summary of included trials is shown in Table 1.

### Placebo-controlled trials

Twenty-two placebo-controlled trials (27 articles) provided data for at least one prespecified outcome measure (Rogers and Friedhoff 1996; Agid et al 1998; Corey-Bloom et al 1998; Rogers et al 1998a, 1998b; Burns et al 1999; Rosler et al 1999; Homma et al 2000; Raskind et al 2000; Tariot et al 2000, 2001; Wilcock et al 2000; Feldman et al 2001; Mohs et al 2001; Rockwood et al 2001, 2006; Wilkinson and Murray 2001; Winblad et al 2001, 2006; Courtney et al 2004; Seltzer et al 2004; Brodaty et al 2005): 14 on cognition; 14 on function; 7 on behavior; and 13 on global assessment of change.

### Cognition

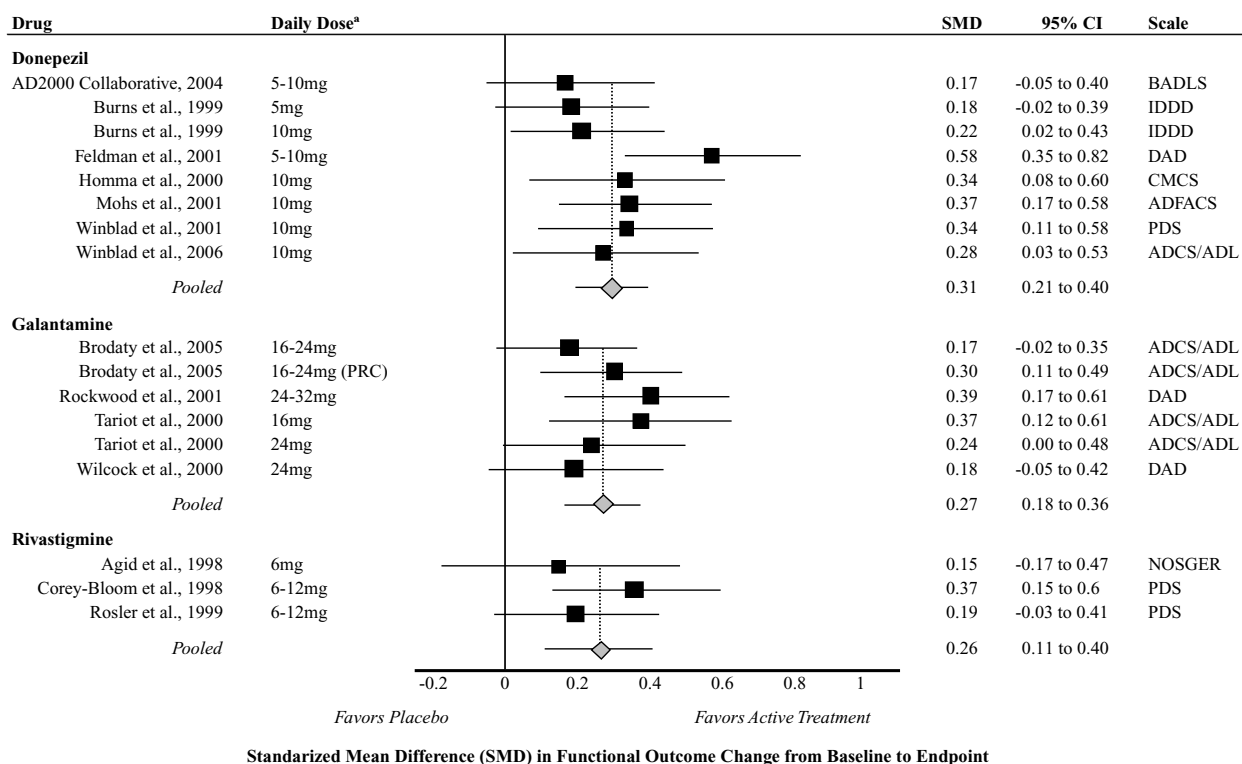
Fourteen studies measured and reported the mean change in ADAS-cog score from baseline to endpoint for active treatment compared with placebo; five on donepezil (Rogers and Friedhoff 1996; Rogers et al 1998a, 1998b; Homma et al 2000; Seltzer et al 2004); seven on galantamine (Raskind et al 2000; Tariot et al 2000; Wilcock et al 2000; Rockwood et al 2001, 2006; Wilkinson and Murray 2001; Brodaty et al 2005); and two on rivastigmine (Corey-Bloom et al 1998; Rosler et al 1999). All of these studies lasted 3 to 6 months and included participants with mild to moderate dementia (except for one which included only participants with mild dementia; see Seltzer et al 2004). Across studies, the average age of participants was 74 years (range 69 to 78 years), and 62% were female (range 50% to 69% female). Limiting these studies to doses recommended in the manufacturers labeling (Figure 1), the pooled weighted mean difference in change between active treatment and placebo was  $-2.67$  (95% confidence interval [CI]  $-3.28$  to  $-2.06$ ) for donepezil,  $-2.76$  (95% CI  $-3.17$  to  $-2.34$ ) for galantamine, and  $-3.01$  (95% CI  $-3.80$  to  $-2.21$ ) for rivastigmine. The  $I^2$  statistic—which reflects the degree of heterogeneity among

**Table 1** Summary of included studies

Drug(s)	Study	Daily dose <sup>a</sup>	Duration (weeks)	Baseline MMSE	Outcome <sup>b</sup>				Quality rating
					Cognition	Function	Behavior	Global	
Donepezil <sup>(D)</sup> vs. Galantamine <sup>(G)</sup>	Jones et al 2004	<sup>(D)</sup> 5–10 mg <sup>(G)</sup> 8–24 mg	12	10–24	ADAS-cog	DAD	–	–	NA <sup>c</sup>
	Wilcock et al 2003	<sup>(D)</sup> 5–10 mg <sup>(G)</sup> 16–24 mg	52	9–18	ADAS-cog	BADLS	NPI	–	NA <sup>c</sup>
Donepezil <sup>(D)</sup> vs. Rivastigmine <sup>(R)</sup>	Bullock et al 2005	<sup>(D)</sup> 5–10 mg <sup>(R)</sup> 3–12 mg	104	10–20	–	ADCS/ADL	NPI	–	Good
	Wilkinson et al 2002	<sup>(D)</sup> 5–10 mg <sup>(R)</sup> 6–12 mg	12	10–26	ADAS-cog	–	–	–	NA <sup>c</sup>
Donepezil	AD2000 Collaborative, 2004	5–10 mg	60	10–27	–	BADLS	NPI	–	Fair
	Burns et al 1999	5–10 mg <sup>d</sup>	24	10–26	–	IDDD	–	CIBIC+	Fair
	Feldman et al 2001	5–10 mg	24	5–17	–	DAD	NPI	–	Good
	Homma et al 2000	5 mg	24	10–26	ADAS-Jcog	CMCS	–	CGI-C	Fair
	Mohs et al 2001	10 mg	54	12–20	–	ADFACS	–	–	Fair
	Rogers et al 1998a	5–10 mg <sup>d</sup>	12	10–26	ADAS-cog	–	–	CIBIC+	Fair
	Rogers et al 1998b	5–10 mg <sup>d</sup>	24	10–26	ADAS-cog	–	–	CIBIC+	Fair
	Rogers and Friedhoff 1996	5 mg <sup>d</sup>	12	10–26	ADAS-cog	–	–	CGI-C	Fair
	Seltzer et al 2004	10 mg	24	21–26	ADAS-cog	–	–	–	Fair
	Tariot et al 2001	5–10 mg	24	5–26	–	–	–	NPI	Fair
	Winblad et al 2001	10 mg	52	10–26	–	PDS	–	–	Fair
Galantamine	Winblad et al 2006	10 mg	24	1–10	–	ADCS/ADL	NPI	CGI-C	Fair
	Brodaty et al 2005 <sup>e</sup>	16–24 mg	26	10–24	ADAS-cog	ADCS/ADL	NPI	CIBIC+	Good
	Raskind et al 2000	24 mg <sup>d</sup>	24	11–24	ADAS-cog	–	–	CIBIC+	Fair
	Rockwood et al 2001	24–32 mg	12	11–24	ADAS-cog	DAD	–	CIBIC+	Fair
	Rockwood et al 2006	24 mg	16	10–25	ADAS-cog	–	NPI	–	Fair
	Tariot et al 2000	16–24 mg <sup>d</sup>	20	10–22	ADAS-cog	ADCS/ADL	NPI	–	Fair
	Wilcock et al 2000	24 mg	24	11–24	ADAS-cog	DAD	–	CIBIC+	Good
	Wilkinson and Murray 2001	18–24 mg <sup>d</sup>	12	13–24	ADAS-cog	–	–	CGI-C	Fair
Rivastigmine	Agid et al 1998	6 mg	13	NR <sup>f</sup>	–	NOSGER	–	–	Fair
	Corey-Bloom et al 1998	6–12 mg	26	10–26	ADAS-cog	PDS	–	CIBIC+	Fair
	Rosler et al 1999	6–12 mg	26	10–26	ADAS-cog	PDS	–	CIBIC+	Fair

**Notes:** <sup>a</sup>Fixed doses outside of the range of the manufacturer label excluded; flexible dosing studies might allow doses outside of labeled range; <sup>b</sup>Does not indicate whether the outcome was primary or secondary in the trial design, and some outcomes were not reported or abstracted; <sup>c</sup>Not applicable; open-label trials were not rated for quality; <sup>d</sup>Value reflects fixed dosing design (compared with flexible dosing); <sup>e</sup>Galantamine and Galantamine Prolonged-Release Capsule (PRC); <sup>f</sup>DSM-III-R mild-to-moderate dementia. **Abbreviations:** MMSE, Mini Mental State Exam; ADAS-cog, Alzheimer's Disease Assessment Scale-Cognitive section; ADAS-Jcog, Japanese translation of ADAS-cog; SIB, Severe Impairment Battery; CGI-C, Clinical Global Impression of Change; ADCS/ADL, Alzheimer's Disease Cooperative Studies Activities of Daily Living Inventory; ADFACS, Alzheimer's Disease Functional Assessment and Change Scale; BADLS, Bristol Activities of Daily Living Scale; CMCS, Caregiver-rated Modified Crichton Scale; DAD, Disability Assessment for Dementia; IDDD, Interview for Deterioration in Daily living activities in Dementia; NOSGER-IADL, Nurses Observation Scale for Geriatric Patients Activities of Daily Living; PDS, Progressive Deterioration Scale; CIBIC+, Clinician Interview-Based Impression of Change Incorporating Caregiver Information.





**Figure 2** Meta-analysis of functional outcomes for active treatment compared with placebo.

**Notes:** \*Limited to doses recommended by product labeling.

**Abbreviations:** SMD, Standardized Mean Difference; PRC, Prolonged Release Capsule; CI, Confidence Interval; ADCS/ADL, Alzheimer's Disease Cooperative Studies Activities of Daily Living Inventory; ADFACS, Alzheimer's Disease Functional Assessment and Change Scale; BADLS, Bristol Activities of Daily Living Scale; CMCS, Caregiver-rated Modified Crichton Scale; DAD, Disability Assessment for Dementia; IDDD, Interview for Deterioration in Daily living activities in Dementia; NOSGER-IADL, Nurses Observation Scale for Geriatric Patients Activities of Daily Living; PDS, Progressive Deterioration Scale.

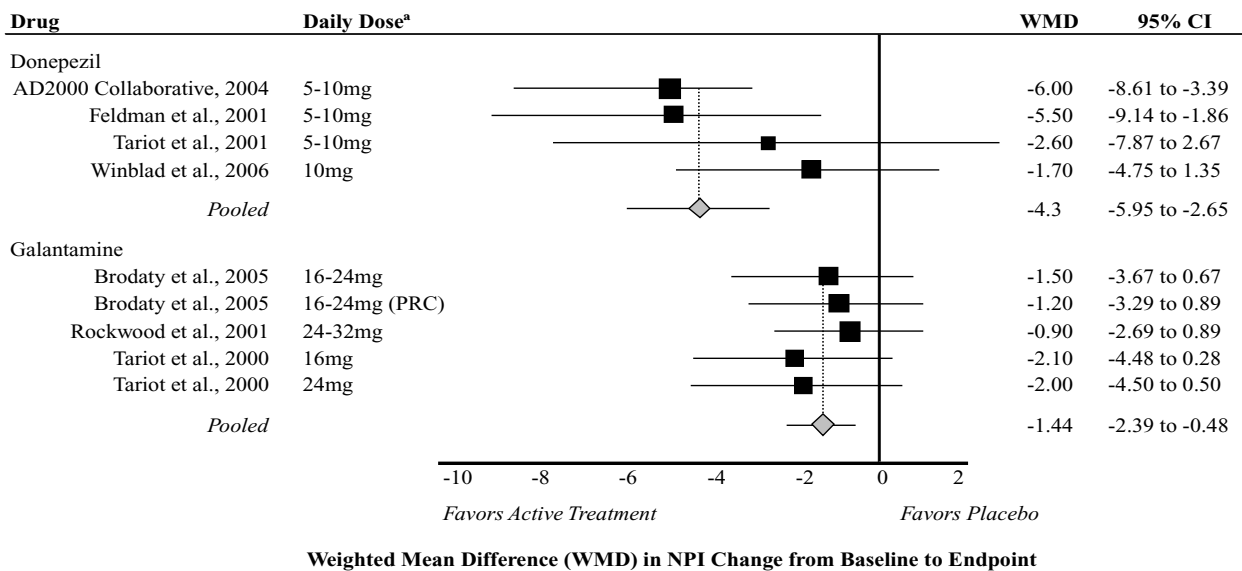
et al 2001; Brodaty et al 2005); and none on rivastigmine. The pooled weighted mean difference in NPI score between active treatment and placebo was  $-4.3$  (95% CI  $-5.95$  to  $-2.65$ ) for donepezil and  $-1.44$  (95% CI  $-2.39$  to  $-0.48$ ) for galantamine (Figure 3). Heterogeneity was moderate among pooled donepezil studies ( $I^2 = 43\%$ ) and low among pooled galantamine studies ( $I^2 = 0\%$ ). The moderate heterogeneity detected among donepezil studies likely was influenced by inclusion of the study by Winblad and colleagues (2006), which was limited to severe dementia. No significant publication bias was detected, and dose stratified analyses did not significantly change overall conclusions (data not shown).

### Global assessment of change

Nine studies reported the number of global responders ( $<4$ ) using the CIBIC+ structured interview; three on donepezil (Rogers et al 1998a, 1998b; Burns et al 1999); four on galantamine (Raskind et al 2000; Wilcock et al 2000; Rockwood et al 2001; Brodaty et al 2005); and two on rivastigmine (Corey-Bloom et al 1998; Rosler et al 1999). These studies lasted 3 to 6 months and included

participants with mild to moderate dementia (mean age 74 years, 63% female). The pooled relative risk of responding for active treatment compared with placebo (Figure 4) was 1.88 (95% CI 1.50 to 2.34) for donepezil, 1.15 (95% CI 0.96 to 1.39) for galantamine, and 1.64 (95% CI 1.29 to 2.09) for rivastigmine. Heterogeneity was low among all pooled analyses ( $I^2 = 0\%$ ). Funnel plots illustrated potential publication bias.

An additional four studies reported the number of global responders ( $<4$ ) using the CGI-C and were included in a sensitivity analysis (Homma et al 2000; Rogers and Friedhoff 1996; Wilkinson and Murray 2001; Winblad et al 2006). These studies also lasted 3 to 6 months and compared with participants in trials measuring the CIBIC+, participants were similar with regard to baseline dementia severity, age, and gender. In sensitivity analyses, combining data for the CIBIC+ and the CGI-C did not significantly influence the pooled estimates for donepezil or rivastigmine, but improved the pooled relative risk estimate for galantamine (RR = 1.21; 95% CI 1.02 to 1.43). Additionally, combining fixed doses to represent the overall number of active treatment responders for a given study did not alter conclusions.



**Figure 3** Meta-analysis of behavior outcomes for active treatment compared with placebo.

**Notes:** <sup>a</sup>Limited to doses recommended by product labeling.

**Abbreviations:** WMD, Weighted Mean Difference; PRC, Prolonged Release Capsule; CI, Confidence Interval; NPI, Neuropsychiatric Inventory.

## Comparative evidence:

### Direct comparisons

Two trials directly compared donepezil with galantamine (Wilcock et al 2003; Jones et al 2004), and two trials (4 articles) directly compared donepezil with rivastigmine (Wilkinson et al 2002; Bullock et al 2005, 2006; Touchon et al 2006). Only one of the four comparative trials was double-blinded (Bullock et al 2005). Relevant outcome data are shown in Figure 5.

### Donepezil vs. galantamine

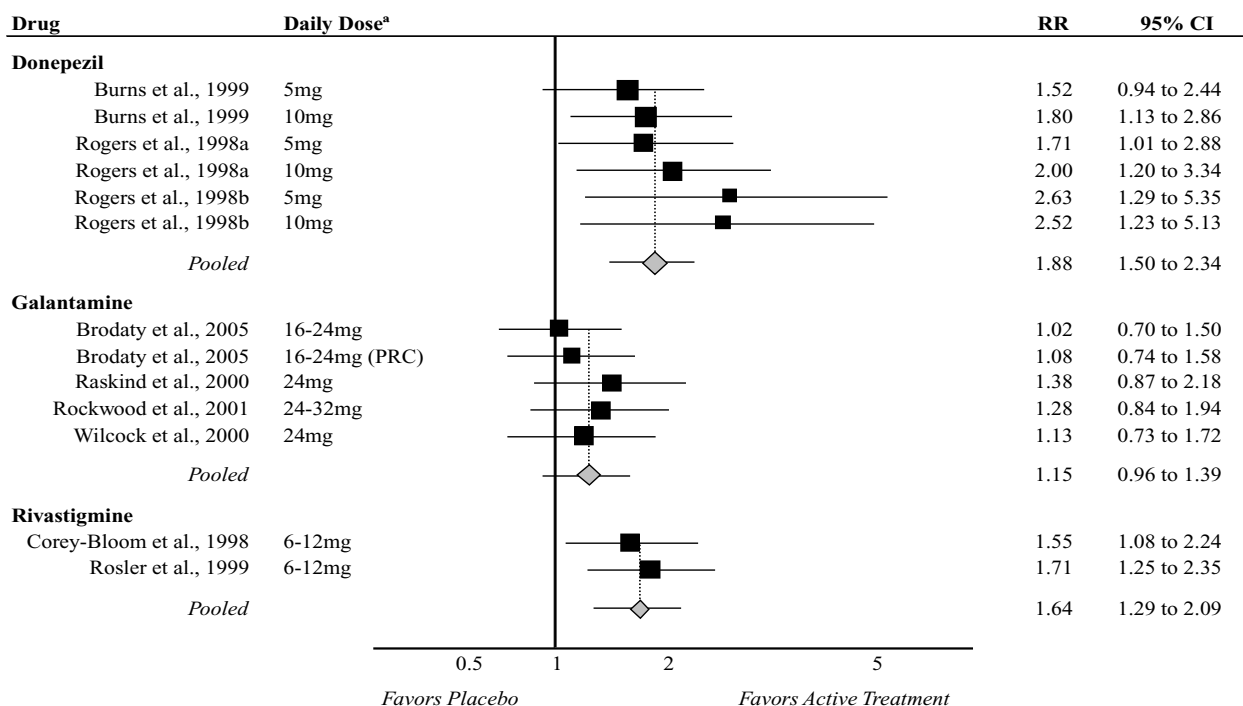
Conflicting head-to-head evidence about the comparative efficacy of donepezil and galantamine comes from two open-label trials; one 52-week trial (Wilcock et al 2003) and one 12-week trial (Jones et al 2004). The 52-week trial compared donepezil 10 mg/day to galantamine 24 mg/day in 182 patients with mild to moderate dementia (Wilcock et al 2003). Relevant outcome measures included the ADAS-cog (cognition), the BADLS (function), and the NPI (behavior). At endpoint, no statistically significant differences between donepezil- and galantamine-treated participants were observed for cognition (ADAS-cog mean change -3.4 vs. -2.2, respectively), function (BADLS mean change 2.7 vs. 2.5, respectively), and behavior (values not reported). In contrast, a shorter 12-week trial compared flexible doses of donepezil 5–10 mg/day (once daily) and galantamine 8–24 mg/day (twice daily) in 120 patients with mild to moderate dementia (Jones et al 2004) and found statistically significant differences in cognition (ADAS-cog mean change -4.7 vs. -2.3,

respectively) and function (DAD mean change 1.6 vs. -0.4), favoring donepezil ( $P < 0.05$ ). The 12- and 52-week studies were both open-label, compromising their validity. Both trials compared relatively equivalent drug doses. However, participants in the 12-week study had less severe baseline MMSE scores than participants in the 52-week trial (mean baseline MMSE = 18 vs. 15, respectively). The 12-week trial was funded by the makers of donepezil, while the 52-week trial was funded by the makers of galantamine.

### Donepezil vs. rivastigmine

Head-to-head evidence for the comparative efficacy of donepezil and rivastigmine also is limited to two trials, with similarly conflicting results as the evidence for donepezil and galantamine. The strongest evidence comes from a good-quality 2-year double-blinded randomized trial (Bullock et al 2005) that compared flexible doses of donepezil (5–10 mg/day) with flexible doses of rivastigmine (3–12 mg/day) in 994 participants with moderate to moderately-severe dementia. Donepezil- and rivastigmine-treated participants had similar changes in cognition (Severe Impairment Battery [SIB] mean change -9.9 vs. -9.3, respectively;  $P > 0.05$ ) and behavior (NPI mean change 2.4 vs. 2.9, respectively;  $P > 0.05$ ) over a 2-year period (Bullock et al 2005). However, rivastigmine-treated participants had statistically significantly better functional (ADCS-ADL -12.8 vs. -14.9, respectively;  $P < 0.05$ ) and global assessment outcomes (Global Deterioration Scale [GDS] 0.58 vs. 0.69, respectively;  $P = 0.05$ ) than donepezil-treated participants.





**Relative Risk (RR) of Response on the CIBIC+ Global Assessment of Change**

**Figure 4** Meta-analysis of clinical global assessment of change for active treatment compared with placebo.

**Notes:** <sup>a</sup>Limited to doses recommended by product labeling.

**Abbreviations:** RR, Relative Risk; PRC, Prolonged Release Capsule; CI, Confidence Interval; CIBIC+, Clinician Interview-Based Impression of Change Incorporating Caregiver Information.

A shorter 12-week open-label trial (Wilkinson et al 2002) compared flexible doses of donepezil (5–10 mg/day) with flexible doses of rivastigmine (6–12 mg/day) in 111 patients with mild to moderate dementia and found no statistically significant differences in cognition (ADAS-cog mean change –0.9 vs. –1.1, respectively;  $P > 0.05$ ) at 12 weeks (Wilkinson et al 2002). Measures of function and behavior were not included in this shorter trial. Aside from apparent difference in duration of follow-up, the largest distinctions between the 12-week and 2-year trials are the double-blinded design (single- vs. double-blinded, respectively) and differences in baseline severity of dementia (mean baseline MMSE = 21 vs. 15, respectively). The 12-week trial was funded by the makers of donepezil, while the 2-year trial was funded by the makers of rivastigmine.

### Comparative evidence: Adjusted indirect comparisons

Data were sufficient to conduct adjusted indirect comparisons of each drug for cognition (ADAS-cog) and global assessment of change (CIBIC+); data were not sufficient to indirectly compare drugs with regard to function, and only donepezil and galantamine could be indirectly compared with regard to behavior (Figure 5). Adjusted indirect

comparison of ADAS-cog change from baseline to endpoint revealed no statistically significant differences in the pooled weighted mean differences among drugs ( $P > 0.05$  for all comparisons). In other words, the drugs produced effects of similar magnitude when compared with placebo. However, adjusted indirect comparisons detected differences among drugs for behavior and global assessment of change. Behavior deteriorated less for donepezil compared with galantamine ( $P = 0.003$ ); data were insufficient to indirectly compare donepezil with rivastigmine or galantamine with rivastigmine. The relative risk of being classified as a global responder statistically significantly favored donepezil and rivastigmine compared with galantamine (RR = 1.63 [ $P < 0.005$ ] and 1.42 [ $P < 0.05$ ], respectively for comparison with galantamine), but did not statistically significantly differ between donepezil and rivastigmine ( $P = 0.4$ ).

### Safety and tolerability

On average across all included trials, 76% (95% CI 70% to 81%) of participants randomized to active treatment reported at least one adverse event. The most frequently reported adverse events were nausea (overall mean 19%; 95% CI 14% to 24%), vomiting (overall mean 13%; 95% CI 9% to 16%), diarrhea (overall mean 11%; 95% CI 9% to 12%), dizziness (overall

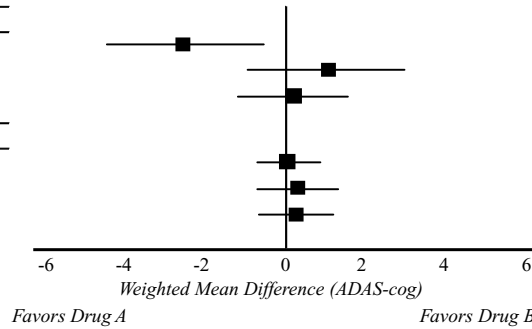
**Cognition (Drug A vs Drug B)**

**Direct Comparison**

Donepezil vs. Galantamine (Jones 2004)  
 Donepezil vs. Galantamine (Wilcock 2003)  
 Donepezil vs. Rivastigmine (Wilkinson 2002)

**Adjusted Indirect Comparison**

Donepezil vs. Galantamine  
 Donepezil vs. Rivastigmine  
 Galantamine vs. Rivastigmine



	WMD	95% CI
Donepezil vs. Galantamine (Jones 2004)	-2.4	-4.47 to -0.33*
Donepezil vs. Galantamine (Wilcock 2003)	1.21	-0.97 to 3.39
Donepezil vs. Rivastigmine (Wilkinson 2002)	0.15	-1.56 to 1.86
Adjusted Indirect Comparison		
Donepezil vs. Galantamine	0.09	-0.65 to 0.83
Donepezil vs. Rivastigmine	0.34	-0.66 to 1.34
Galantamine vs. Rivastigmine	0.25	-0.65 to 1.15

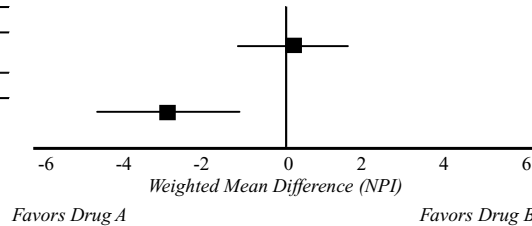
**Behavior (Drug A vs Drug B)**

**Direct Comparison**

Donepezil vs. Rivastigmine (Bullock 2005)

**Adjusted Indirect Comparison**

Donepezil vs. Galantamine

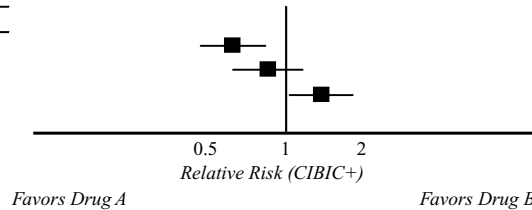


	WMD	95% CI
Donepezil vs. Rivastigmine (Bullock 2005)	0.54	-1.68 to 2.70
Adjusted Indirect Comparison		
Donepezil vs. Galantamine	-2.86	-4.77 to -0.95*

**Global Assessment (Drug A vs Drug B)**

**Adjusted Indirect Comparison**

Donepezil vs. Galantamine  
 Donepezil vs. Rivastigmine  
 Galantamine vs. Rivastigmine



	RR	95% CI
Donepezil vs. Galantamine	0.61	0.49 to 0.82*
Donepezil vs. Rivastigmine	0.87	0.63 to 1.21
Galantamine vs. Rivastigmine	1.42	1.05 to 1.93*

**Figure 5** Comparative evidence for donepezil, galantamine, and rivastigmine<sup>a</sup>.

**Notes:** <sup>a</sup>Limited to comparisons with sufficient data for a single outcome measure and to doses recommended by product labeling.

**Abbreviations:** WMD, Weighted Mean Difference (reflects the pooled difference for Drug A – Drug B); RR, Relative Risk (reflects the relative risk of responding with Drug B/Drug A); CI, Confidence Interval; ADAS-cog, Alzheimer’s Disease Assessment Scale-Cognitive section; NPI, Neuropsychiatric Inventory; CIBIC+, Clinician Interview-Based Impression of Change Incorporating Caregiver Information.

mean 10%; 95% CI 8% to 12%), and weight loss (overall mean 9%; 95% CI 6% to 11%). With the exception of diarrhea (mean frequency: donepezil 12%; galantamine 8%; rivastigmine 13%), the mean frequency of these events was consistently lowest for donepezil and highest for rivastigmine (nausea 11%, 24%, and 44%; vomiting 7%, 14%, and 30%; dizziness 8%, 10%, and 22%; and weight loss 7%, 10%, and 11%, respectively for donepezil, galantamine, and rivastigmine).

Overall, 26% (95% CI 21% to 31%) of participants randomized to active treatment withdrew from trials, approximately half of which withdrew specifically because of adverse events (overall mean 13%; 95% CI 10% to 16%). The frequency of withdrawals and withdrawals due to adverse events also was lowest among donepezil trials and highest among rivastigmine trials. Withdrawals and withdrawals due to adverse events were 24% (95% CI 16% to 32%) and 11% (95% CI 8% to 14%), respectively, for donepezil; 27% (95% CI 21% to 33%) and 14% (95% CI 10% to 18%), respectively, for galantamine; and 28% (95% CI 15% to 40%) and 21% (95% CI 12% to 31%),

respectively, for rivastigmine. In our meta-analysis of placebo-controlled trials, the pooled relative risk of withdrawing for any reason was 1.1 (95% CI 0.9 to 1.3) for donepezil, 1.6 (95% CI 1.3 to 1.9) for galantamine, and 2.3 (95% CI 1.8 to 2.9) for rivastigmine. Similarly, the pooled relative risk of withdrawing because of adverse events was 1.3 (95% CI 0.9 to 1.8) for donepezil, 2.0 (95% CI 1.4 to 2.8) for galantamine, and 3.6 (95% CI 2.6 to 5.1) for rivastigmine. These analyses included all placebo-controlled studies included in our assessment of efficacy (Table 1), except for the AD2000 Collaborative Group study (2004) which did not report sufficient data. Heterogeneity was moderate for analyses of donepezil and galantamine ( $I^2$  between 40% to 50% for all), but low ( $I^2 = 0%$ ) for analyses of rivastigmine. Factors such as drug dose and baseline disease severity varied among studies, and likely contributed to heterogeneity.

**Discussion**

Meta-analyses of placebo-controlled data support the drugs’ modest overall benefits for stabilizing or slowing decline in

cognition, function, behavior, and clinical global change. Evidence directly comparing one drug with another is limited to four trials, three of which used an open-label design. Of two open-label trials comparing donepezil with galantamine (Wilcock et al 2003; Jones et al 2004), one found no statistically significant differences in efficacy (Wilcock et al 2003), while one found statistically significantly better cognition and function outcomes for donepezil (Jones et al 2004). One open-label trial (Wilkinson et al 2002) and one double-blinded trial (Bullock et al 2005) directly compared donepezil with rivastigmine. Both trials found drugs to be similar with regard to cognitive outcomes, although the double-blinded study reported small but statistically significant differences in function favoring rivastigmine compared with donepezil. Adjusted indirect comparisons found drugs to be similar with regard to cognitive outcomes. However, donepezil performed statistically significantly better than galantamine with regard to behavior, and both donepezil and rivastigmine performed statistically significantly better than galantamine with regard to global assessment.

Results of our adjusted indirect comparisons are consistent with findings of some head-to-head trials, but conflict with results of other comparative studies. For example, our indirect comparison of cognitive outcomes did not reveal statistically significant differences among drugs—a conclusion similar to most comparative trials (Wilkinson et al 2002; Wilcock et al 2003; Bullock et al 2005) and a meta-analysis by Harry and Zakzanis (2005). However, Jones and colleagues (2004) reported greater improvements in cognition for donepezil- compared with galantamine-treated patients, a finding inconsistent with other evidence. Interestingly, our adjusted indirect comparisons paralleled the direction of the findings of Jones and colleagues for other outcome measures, even though the Jones study (2004) did not measure these outcomes. For instance, our adjusted indirect comparison favored donepezil over galantamine for measures of behavior (NPI) and global assessment of change (CIBIC+). No other evidence directly comparing donepezil and galantamine on these outcome measures are available to contrast this finding. Although our adjusted indirect comparison found donepezil and rivastigmine to be similar with regard to clinical global assessment on the CIBIC+, a good-rated comparative trial found modest differences ( $P = 0.05$ ) favoring rivastigmine over donepezil. However, the comparative trial was conducted in patients with moderate to severe dementia and used the GDS rather than the CIBIC+ to assess global change. Thus, differences in measurement scale and trial population confound this comparison.

The most common adverse events reported in trials were nausea, vomiting, diarrhea, dizziness, and weight loss. Across studies, the frequency in which these events were reported was generally lowest for donepezil and highest for rivastigmine. This trend paralleled overall withdrawal rates and withdrawals due to adverse events. The relative risk of withdrawing for any reason or because of adverse events was similar for donepezil compared with placebo, but the relative risk was statistically significantly greater for galantamine and rivastigmine compared with placebo.

Although the frequency in which adverse events were reported and analysis of withdrawal rates provide a compelling argument in favor of donepezil with regard to tolerability, heterogeneity in these data must be considered. First, studies differ in how adverse events are assessed and reported. Most studies did not specify adverse a priori, and reporting of specific events varied (eg, report all events with incidence  $>5\%$  vs. report events statistically significantly different from placebo). Second, the frequency of specific events varied within individual studies and across studies for a given drug. Within study variance could be explained in part by differences in doses, with higher adverse event rates generally reported among higher doses (Ritchie et al 2004). In some cases, differences in event rates could be explained by differences in formulation. For example, one study of galantamine compared the immediate release and the extended release formulation with placebo (Brodaty et al 2005). A post hoc comparison of these formulations (Dunbar et al 2006) illustrated that patients randomized to the extended release formulation had statistically significantly fewer days with nausea than participants randomized to the immediate release formulation (18% vs. 38%;  $P = 0.014$ ).

Many different measurement scales are used in assessing outcomes of Alzheimer's treatment. We chose to focus on four overall outcome domains: cognition, function, behavior, and clinical global assessment of change. Within these general domains, we limited our data abstraction to specific measures that were commonly used across trials. For example, we abstracted data only for the ADAS-cog scale for cognition. Although this is a relatively common scale used to assess cognition in trials of mild to moderate dementia, its use in patients with more severe dementia is subject to floor effects (Schmitt et al 2006). Measurement scales have been developed for use in patients with more severe dementia (eg, the SIB), but the number of trials conducted in patients with severe dementia and using these scales were too few for us to pool data. We chose to exclude studies that did not use predefined outcome measures, thus indirectly limiting our

analysis to populations with mild to moderate dementia—at least for some outcome domains. For function, because measurement scales were so extensively varied, we used a standardized effect size analysis. This analysis included measures believed to be sensitive in more severe disease (eg, the ADCS-ADL [Galasko et al 1997] and DAD [Gelinas et al 1999]), but it is subject to other limitations such as interpretation of meaningful differences (Cohen 1988). As additional evidence accrues for specific measurement scales, additional meta-analyses should test the sensitivity of our findings among patients with more severe disease.

A number of other factors limit the conclusions of our analysis. First, Alzheimer's disease is progressive, and patients decline at different rates. This may have implications when pooling data across studies. For example, one study conducted in patients with early-stage Alzheimer's disease illustrated little or no decline in cognition among the placebo-treated participants (Seltzer et al 2004), while a second study reported nearly a 2-point decline in cognition (ADAS-cog) among placebo-treated participants with mild-to-moderate dementia (Rogers et al 1998b). In our analysis, we pooled data from all studies regardless of dementia severity, potentially biasing our results. Second, although we limited our review to doses within the manufacturers' current recommendations, we still included a range of fixed and flexible doses. A meta-analysis by Ritchie and colleagues (2004) demonstrated the dose-response relationship for these drugs by pooling studies for specific doses. Although we present only the overall analysis for each outcome measure, we also conducted dose-stratified sensitivity analyses. Stratifying by dose illustrated a dose-response relationship, but did not change conclusions of individual meta-analyses or indirect comparisons. Other population inclusion and exclusion criteria might also influence our results. Although most trials used accepted methods for confirming the diagnosis of Alzheimer's (eg, diagnosis consistent with the DSM-IV and the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria), studies did not consistently include or exclude patients who also had symptoms suggestive of concomitant Lewy body disease, patients with other co-morbid diagnoses, or patients using other medications. The implications of these distinctions may be significant. For example, a large prospective cohort study found that patients without concomitant disease at baseline had a 2-fold greater likelihood of being classified as a cognitive responder at 9 months (Raschetti et al 2005). Arguably, however, factors such as co-morbid illness and

variations in other medication use are representative of the environment for treating Alzheimer's disease. Still, in the context of meta-analysis, variation in patient populations and trial design can bias conclusions. This potential concern is likely reflected by the moderate to high heterogeneity we detected in some meta-analyses.

Finally, our analysis was limited to studies identified at the time of our literature search (ie, July 2007). New evidence continues to accrue and should be considered in future reviews. For example, a recent trial (Howard et al 2007) compared donepezil to placebo in patients with Alzheimer's disease who had clinically significant agitation. Donepezil was not more effective than placebo in treating agitation or other behavioral symptoms, even though cognitive measures showed modest benefit from donepezil compared with placebo. If included, these findings may have influenced conclusions of our adjusted indirect comparisons.

## Conclusions

Compared with placebo, the cholinesterase inhibitors donepezil, galantamine, and rivastigmine are able to stabilize or slow decline in cognition, function, behavior, and global change. No clear evidence exists to determine whether one of these drugs is more efficacious than another, although adjusted indirect comparisons suggest that donepezil and rivastigmine may be slightly more efficacious than galantamine, at least as reflected by some outcome measures. The incidence of common adverse events appears to be lowest with donepezil and highest with rivastigmine. Additional high quality comparative evidence is needed to confirm these conclusions.

## Disclosure

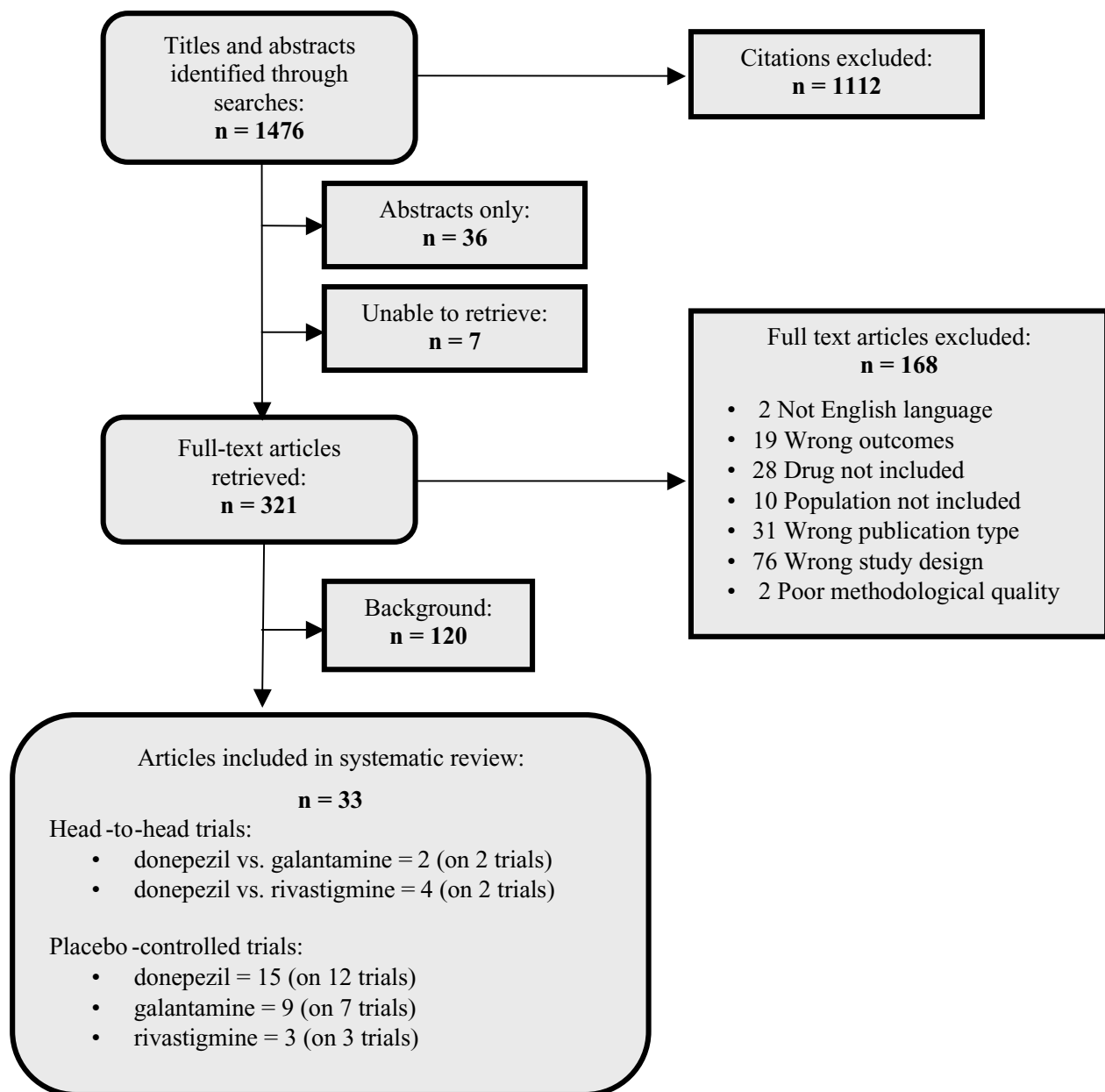
The work was partially funded by the Cecil G Sheps Center for Health Services Research through a sub-contract with the Center for Evidence-Based Policy; Oregon Health and Science University. Dr. Hansen is supported by grant K12RR023248.

## References

- Agid Y, Dubois B, Anand R, et al. 1998. Efficacy and tolerability of rivastigmine in patients with dementia of the Alzheimer type. *Curr Ther Res Clin Exp*, 59:837–45.
- American Psychiatric Association. 2007. Practice guideline for the treatment of patients with Alzheimer's disease and other dementias, 2nd Ed [online]. Accessed February 19, 2008. URL: <http://www.psychiatryonline.com/pracGuide/loadGuidelinePdf.aspx?file=AlzPG101007>.
- Beier MT. 2007. Treatment strategies for the behavioral symptoms of Alzheimer's disease: focus on early pharmacologic intervention. *Pharmacotherapy*, 27:399–411.
- Birks J. 2006. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst Rev*, 1:CD005593.

- Brodaty H, Corey-Bloom J, Potocnik FC, et al. 2005. Galantamine prolonged-release formulation in the treatment of mild to moderate Alzheimer's disease. *Dement Geriatr Cogn Disord*, 20:120–32.
- Bucher HC, Guyatt GH, Griffith LE, et al. 1997. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*, 50:683–91.
- Bullock R, Bergman H, Touchon J, et al. 2006. Effect of age on response to rivastigmine or donepezil in patients with Alzheimer's disease. *Curr Med Res Opin*, 22:483–94.
- Bullock R, Touchon J, Bergman H, et al. 2005. Rivastigmine and donepezil treatment in moderate to moderately-severe Alzheimer's disease over a 2-year period. *Curr Med Res Opin*, 21:1317–27.
- Burns A, Rossor M, Hecker J, et al. 1999. The effects of donepezil in Alzheimer's disease - results from a multinational trial. *Dement Geriatr Cogn Disord*, 10:237–44.
- Centre for Reviews and Dissemination. 2001. Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. National Health Service, York, United Kingdom.
- Cohen J. 1988. Statistical power for the behavioral sciences. 2nd ed., Erlbaum, Hillsdale, NJ.
- Corey-Bloom JR, Anand JV, Veach J, et al. 1998. A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor, in patients with mild to moderately severe Alzheimer's disease. *Int J Geriatr Psychopharmacol*, 1:55–65.
- Courtney C, Farrell D, Gray R, et al. 2004. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet*, 363:2105–15.
- Cummings JL, Frank JC, Cherry D, et al. 2002. Guidelines for managing Alzheimer's disease: Part II. Treatment. *Am Fam Physician*, 65:2525–34.
- Cummings JL, Mega M, Gray K, et al. 1994. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44:2308–14.
- Dunbar F, Zhu Y, Brashear HR. 2006. Post hoc comparison of daily rates of nausea and vomiting with once- and twice-daily galantamine from a double-blind, placebo-controlled, parallel-group, 6-month study. *Clin Ther*, 28:365–72.
- Egger M, Davey Smith G, Schneider M, et al. 1997. Bias in meta-analysis detected by a simple, graphical test. *Bmj*, 315:629–34.
- Feldman H, Gauthier S, Hecker J, et al. 2001. A 24-week, randomized, double-blind study of donepezil in moderate to severe Alzheimer's disease. *Neurology*, 57:613–20.
- Forchetti CM. 2005. Treating patients with moderate to severe Alzheimer's disease: implications of recent pharmacologic studies. *Prim Care Companion J Clin Psychiatry*, 7:155–61.
- Galasko D, Bennett D, Sano M, et al. 1997. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord*, 11(Suppl 2): S33–9.
- Geldmacher DS. 2003. Long-term cholinesterase inhibitor therapy for Alzheimer's disease: Practical considerations for the primary care physician. *Prim Care Companion J Clin Psychiatry*, 5:251–9.
- Geldmacher DS. 2007. Treatment guidelines for Alzheimer's disease: redefining perceptions in primary care. *Prim Care Companion J Clin Psychiatry*, 9:113–21.
- Gelinas I, Gauthier L, McIntyre M, et al. 1999. Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. *Am J Occup Ther*, 53:471–81.
- Glenny AM, Altman DG, Song F, et al. 2005. Indirect comparisons of competing interventions. *Health Technol Assess*, 9:1–134, iii–iv.
- Hansen RA, Gartlehner G, Lohr KN, et al. 2007. Functional outcomes of drug treatment in Alzheimer's disease: A systematic review and meta-analysis. *Drugs Aging*, 24:155–67.
- Harris RP, Helfand M, Woolf SH, et al. 2001. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*, 20:21–35.
- Harry RD, Zakzanis KK. 2005. A comparison of donepezil and galantamine in the treatment of cognitive symptoms of Alzheimer's disease: a meta-analysis. *Hum Psychopharmacol*, 20:183–7.
- Hebert LE, Scherr PA, Bienias JL, et al. 2003. Alzheimer's disease in the US population: prevalence estimates using the 2000 Census. *Arch Neurol*, 60:1119–22.
- Higgins JP, Thompson SG, Deeks JJ, et al. 2003. Measuring inconsistency in meta-analyses. *BMJ*, 327:557–60.
- Homma A, Takeda M, Imai Y, et al. 2000. Clinical efficacy and safety of donepezil on cognitive and global function in patients with Alzheimer's disease. A 24-week, multicenter, double-blind, placebo-controlled study in Japan. E2020 Study Group. *Dement Geriatr Cogn Disord*, 11:299–313.
- Howard RJ, Juszcak E, Ballard CG, et al. 2003. Donepezil for the treatment of agitation in Alzheimer's disease. *N Engl J Med*, 357:1382–92.
- Jones RW, Soininen H, Hager K, et al. 2004. A multinational, randomised, 12-week study comparing the effects of donepezil and galantamine in patients with mild to moderate Alzheimer's disease. *Int J Geriatr Psychiatry*, 19:58–67.
- Kaduszkiewicz H, Zimmermann T, Beck-Bornholdt HP, et al. 2005. Cholinesterase inhibitors for patients with Alzheimer's disease: systematic review of randomised clinical trials. *BMJ*, 331:321–7.
- Knopman DS, Knapp MJ, Gracon SI, et al. 1994. The Clinician Interview-Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. *Neurology*, 44:2315–21.
- Lancot KL, Herrmann N, LouLou MM. 2003a. Correlates of response to acetylcholinesterase inhibitor therapy in Alzheimer's disease. *J Psychiatry Neurosci*, 28:13–26.
- Lancot KL, Herrmann N, Yau KK, et al. 2003b. Efficacy and safety of cholinesterase inhibitors in Alzheimer's disease: a meta-analysis. *CMAJ*, 169:557–64.
- Loveman E, Green C, Kirby J, et al. 2006. The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. *Health Technol Assess*, 10, iii–iv, ix–xi, 1–160.
- Loy C, Schneider L. 2006. Galantamine for Alzheimer's disease and mild cognitive impairment. *Cochrane Database Syst Rev*, 1:CD001747.
- Masterman D. 2004. Cholinesterase inhibitors in the treatment of Alzheimer's disease and related dementias. *Clin Geriatr Med*, 20:59–68.
- Mohs RC, Doody RS, Morris JC, et al. 2001. A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. *Neurology*, 57:481–8.
- [NICE] National Institute for Health and Clinical Excellence. 2007. Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease (amended): NICE technology appraisal guidance 111 (amended). National Health Service, London, England.
- Raschetti R, Maggini M, Sorrentino GC, et al. 2005. A cohort study of effectiveness of acetylcholinesterase inhibitors in Alzheimer's disease. *Eur J Clin Pharmacol*, 61:361–8.
- Raskind MA, Peskind ER, Wessel T, et al. 2000. Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology*, 54:2261–8.
- Ritchie CW, Ames D, Clayton T, et al. 2004. Metaanalysis of randomized trials of the efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer disease. *Am J Geriatr Psychiatry*, 12:358–69.
- Rockwood K, Fay S, Song X, et al. 2006. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *Cmaj*, 174:1099–105.
- Rockwood K, Mintzer J, Truyen L, et al. 2001. Effects of a flexible galantamine dose in Alzheimer's disease: a randomised, controlled trial. *J Neurol Neurosurg Psychiatry*, 71:589–95.
- Rogers SL, Doody RS, Mohs RC, et al. 1998a. Donepezil improves cognition and global function in Alzheimer disease: a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. *Arch Intern Med*, 158:1021–31.
- Rogers SL, Farlow MR, Doody RS, et al. 1998b. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology*, 50:136–45.

- Rogers SL, Friedhoff LT. 1996. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US Multicentre, Randomized, Double-Blind, Placebo-Controlled Trial. The Donepezil Study Group. *Dementia*, 7:293–303.
- Rosen WG, Mohs RC, Davis KL. 1984. A new rating scale for Alzheimer's disease. *Am J Psychiatry*, 141:1356–64.
- Rosler M, Anand R, Cicin-Sain A, et al. 1999. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial. *BMJ*, 318:633–8.
- Schmitt FA, van Dyck CH, Wichems CH, et al. 2006. Cognitive response to memantine in moderate to severe Alzheimer disease patients already receiving donepezil: an exploratory reanalysis. *Alzheimer Dis Assoc Disord*, 20:255–62.
- Schneider LS, Olin JT, Doody RS, et al. 1997. Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change. The Alzheimer's Disease Cooperative Study. *Alzheimer Dis Assoc Disord*, 11(Suppl 2):S22–32.
- Seltzer B, Zolnouni P, Nunez M, et al. 2004. Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial. *Arch Neurol*, 61:1852–6.
- Takeda A, Loveman E, Clegg A, et al. 2006. A systematic review of the clinical effectiveness of donepezil, rivastigmine and galantamine on cognition, quality of life and adverse events in Alzheimer's disease. *Int J Geriatr Psychiatry*, 21:17–28.
- Tariot PN, Cummings JL, Katz IR, et al. 2001. 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*, 49:1590–9.
- Tariot PN, Solomon PR, Morris JC, et al. 2000. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology*, 54:2269–76.
- Touchon J, Bergman H, Bullock R, et al. 2006. Response to rivastigmine or donepezil in Alzheimer's patients with symptoms suggestive of concomitant Lewy body pathology. *Curr Med Res Opin*, 22:49–59.
- Trinh NH, Hoblyn J, Mohanty S, et al. 2003. Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: a meta analysis. *JAMA*, 289:210–16.
- Wilcock G, Howe I, Coles H, et al. 2003. A long-term comparison of galantamine and donepezil in the treatment of Alzheimer's disease. *Drugs Aging*, 20:777–89.
- Wilcock GK, Lilienfeld S, Gaens E. 2000. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. Galantamine International-1 Study Group. *BMJ*, 321:1445–9.
- Wilkinson D, Murray J. 2001. Galantamine: a randomized, double-blind, dose comparison in patients with Alzheimer's disease. *Int J Geriatr Psychiatry*, 16:852–7.
- Wilkinson DG, Passmore AP, Bullock R, et al. 2002. A multinational, randomised, 12-week, comparative study of donepezil and rivastigmine in patients with mild to moderate Alzheimer's disease. *Int J Clin Pract*, 56:441–6.
- Winblad B, Engedal K, Soininen H, et al. 2001. A 1-year, randomized, placebo-controlled study of donepezil in patients with mild to moderate AD. *Neurology*, 57:489–95.
- Winblad B, Kilander L, Eriksson S, et al. 2006. Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled study. *Lancet*, 367:1057–65.



Appendix I Literature review flow diagram.

